

**FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO**



**Avaliação de Agentes de Conversação:**  
A influência de elementos Multimédia

**César Filipe Monteiro Brandão**

Mestrado em Multimédia

Orientador: Prof. Doutor Luís Paulo Reis

Coorientador: Prof.<sup>a</sup> Ana Paula Rocha

Outubro de 2012



# **Avaliação de Agentes de Conversação: A influência de elementos Multimédia**

**César Filipe Monteiro Brandão**

Mestrado em Multimédia

Aprovado em provas públicas pelo Júri:

Presidente: Doutor Eurico Manuel Elias Morais Carrapatoso (Professor Auxiliar da Faculdade de Engenharia da Universidade do Porto)

Vogal Externo: Doutor Pedro Manuel Henriques da Cunha (Professor Auxiliar convidado da Faculdade de Ciências e Tecnologia da Universidade de Coimbra)

Orientador: Doutor Luis Paulo Gonçalves dos Reis (Professor Associado da Universidade do Minho)

Coorientador: Doutora Ana Paula Cunha da Rocha (Professora Auxiliar da Faculdade de Engenharia da Universidade do Porto)

---

30 de Outubro de 2012



# Abstract

*In this thesis, the influence of different multimedia elements, when used in a virtual conversation environment, has been studied considering the actual state of technology of Embodied Conversational Agents in Multimedia Applications.*

*Using the “Imitation Game” idealized by Alan Turing, as a starting point and as motivation, a platform has been developed that allows, through a series of tests, to gather information regarding the efficiency of several changes in the multimedia used in the presentation of the Virtual Environment. This way it is possible to identify multimedia characteristics that could be used to increase the efficiency in the use of Embodied Conversational Agents in all kind of applications.*

*Using the developed platform, several user tests were performed, using different configurations and gathered data through the user interaction with the platform and also through a survey.*

# Resumo

Neste trabalho estudou-se a influência de diferentes elementos multimédia em um ambiente virtual de conversação tendo em consideração o estado atual da tecnologia de Agentes de Conversação e de Aplicações Multimédia.

Utilizando o “Jogo da Imitação” idealizado por Alan Turing, como ponto de partida e como motivação, desenvolveu-se uma plataforma que permite, através de um conjunto de testes, recolher informação sobre a eficiência de diferentes alterações nos conteúdos multimédia utilizados na apresentação do Ambiente Virtual. Deste modo é possível identificar características multimédia que possam ser utilizadas para aumentar a eficiência do uso de Agentes de Conversação em todo o tipo de aplicações.

Usando a plataforma desenvolvida, foram realizados testes com vários utilizadores, utilizando diferentes configurações e recolhidos dados através quer da interação do utilizador com a ferramenta de teste quer através de um questionário realizado.



# Agradecimentos

Gostaria de agradecer a todos aqueles que tornaram possível a concretização de este trabalho.

Ao Prof. Doutor Luís Paulo Reis e Prof.<sup>a</sup> Ana Paula Rocha pela orientação e ajuda, sem os quais não teria sido possível a elaboração de este documento.

A todos os que participaram na utilização experimental da ferramenta com um agradecimento especial a Mário Gonçalves e Gil Teixeira





# Índice

<b>1. Introdução</b>	<b>13</b>
1.1 Enquadramento .....	13
1.2 Motivação.....	13
1.3 Objetivos .....	13
1.4 Estrutura do Relatório .....	14
<b>2. Agentes de Conversação</b>	<b>16</b>
2.1 O trabalho de Alan Turing .....	16
2.1.1 A máquina de Turing .....	16
2.1.2 O jogo da Imitação .....	18
2.1.3 O Prémio Loebner .....	19
2.1.3.1 As primeiras edições do Prémio Loebner.....	19
2.1.3.2 O Prémio de Loebner sem restrições.....	21
2.2 Agentes de Conversação .....	22
2.2.1 ELIZA, o Psicoterapeuta Rogeriano.....	22
2.2.2 A.L.I.C.E. ....	23
2.2.3 PARRY, o paciente Paranoico.....	27
2.2.4 Jaberwacky .....	27
2.2.5 Elbot .....	29
2.2.6 Suzette .....	29
2.3 Conclusão.....	30
<b>3. Agentes de Conversação em Ambientes Virtuais</b>	<b>31</b>
3.1 A influência da linguagem e do domínio de conhecimento .....	31
3.1.1 AZ-ALICE e TARA .....	32
3.1.1.1 AZ-ALICE .....	32
3.1.1.2 TARA.....	34
3.1.2 Emile.....	36
3.1.3 Autotutor.....	37
3.2 O efeito Persona .....	39
3.2.1 Herman, o inseto.....	39
3.2.2 Freudbot e Piagetbot.....	40
3.2.3 Outros estudos sobre a influência da representação visual.....	41
3.3 Agentes de Conversação com representação emocional .....	44

---

3.3.1	Representação de emoções .....	45
3.4	Conclusão.....	50
<b>4.</b>	<b>Plataforma desenvolvida</b>	<b>51</b>
4.1	Especificação .....	51
4.1.1	Servidor de Conversação .....	51
4.1.2	Agente de Conversação .....	52
4.1.3	Funcionalidades a Implementar.....	53
4.1.4	Recolha e armazenamento de Informação´.....	54
4.2	Implementação .....	54
4.2.1	Servidor de Conversação .....	54
4.2.2	Agente de Conversação .....	55
4.2.3	Cliente de Conversação .....	56
4.2.3.1	Tipos de sessão.....	56
4.2.4	Recolha e Armazenamento de Informação.....	59
4.3	Conclusão.....	59
<b>5.</b>	<b>Avaliação experimental</b>	<b>61</b>
5.1	Configuração da plataforma.....	61
5.2	Utilização experimental da plataforma .....	62
5.2.1	Interface de conversação simples .....	62
5.2.2	Interface de conversação com Avatares emotivos.....	64
5.2.3	Interface de conversação com Avatares emotivos e imagem de emoção. ....	66
5.3	Conclusão.....	69
<b>6.</b>	<b>Conclusões e Perspetivas de Desenvolvimento</b>	<b>70</b>

# Lista de Figuras

ILUSTRAÇÃO 1. EXEMPLO DE UMA FITA INFINITA E CABEÇA DE LEITURA TAL COMO DESCRITO POR TURING [COOPER 2004].....	17
ILUSTRAÇÃO 2. EXEMPLO DO FORMULÁRIO UTILIZADO PARA CLASSIFICAR OS CONCORRENTES NA PRIMEIRA EDIÇÃO DA COMPETIÇÃO LOEBNER. [SHIEBER STUART 1993] .....	19
ILUSTRAÇÃO 3. INTERFACE DO AGENTE DE CONVERSAÇÃO GEORGE.....	28
ILUSTRAÇÃO 4. APARÊNCIA DO AGENTE HUMANA E DO AGENTE ICÓNICO [XIAO, STASKO E CATRAMBONE 2002] .....	43
ILUSTRAÇÃO 5. EMOÇÕES BÁSICAS (REPUGNÂNCIA, FELICIDADE, TRISTEZA, RAIVA, MEDO, E SURPRESA) [VINCIARELLI, PANTIC E BOURLARD 2009].....	46
ILUSTRAÇÃO 6. EXEMPLO DE CONVERSAÇÃO ENTRE AGENTE E UTILIZADOR .....	52
ILUSTRAÇÃO 7. EXEMPLO DE CONVERSAÇÃO ENTRE DOIS UTILIZADORES.....	53
ILUSTRAÇÃO 8. INTERFACE DE ADMINISTRAÇÃO DE AGENTES PANDORABOTS [PANDORABOTS] .	55
ILUSTRAÇÃO 9. INTERFACE DE AUTENTICAÇÃO OU REGISTO DO UTILIZADOR .....	56
ILUSTRAÇÃO 10. EXEMPLO DO INTERFACE DE CONVERSAÇÃO SIMPLES .....	57
ILUSTRAÇÃO 11 EXEMPLO DO INTERFACE DE CONVERSAÇÃO COM AVATARES EMOTIVOS.....	57
ILUSTRAÇÃO 12 EXEMPLO DO INTERFACE DE CONVERSAÇÃO COM AVATARES E GRÁFICO DE EMOÇÕES .....	58
ILUSTRAÇÃO 13 EXEMPLO DO INTERFACE DE CONVERSAÇÃO COM AVATARES E IMAGEM DE EMOÇÃO.....	58



## Capítulo 1

# 1. Introdução

## 1.1 Enquadramento

O trabalho descrito neste relatório enquadra-se na tecnologia de Agentes de Conversação que são objeto de desenvolvimento e estudo principalmente na área da Inteligência Artificial.

## 1.2 Motivação

A motivação principal para a realização deste trabalho resulta da importância crescente da utilização eficiente de agentes de conversação em diversos contextos, como ferramentas de *E-Learning* ou mesmo a possível automatização de sistemas de atendimento ao cliente. Assim sendo, torna-se essencial identificar e quantificar os problemas que surgem, inevitavelmente, durante a comunicação entre uma pessoa e um agente de conversação de forma a tornar esta experiência mais agradável para o utilizador. Desta forma, pretende-se analisar o impacto de alterações na plataforma de comunicação com o objetivo de melhorar e maximizar a eficiência da utilização do Agente de conversação.

É também um fator de motivação, a utilização de elementos multimédia para complementar o estudo de uma questão tão importante como a levantada por Turing em 1950 com elementos proporcionados pela tecnologia atual.

## 1.3 Objetivos

O objetivo desta dissertação é avaliar o desempenho de uma tecnologia de agentes de conversação, recorrendo ao uso de uma ferramenta que permita simular o teste de Turing. Propõe-se enriquecer o agente de conversação com elementos multimédia que permitam uma visualização do seu estado emocional durante a conversação, com o intuito de fazer o agente parecer o mais “humano” possível.

A avaliação do desempenho é efetuada através da recolha e análise de resultados de testes, recorrendo a uma base de conhecimento específica.

O objetivo de este trabalho não é desenvolver um agente capaz de superar o Teste de Turing, mas sim de desenvolver uma plataforma que permita a avaliação e consequente melhoria da utilização de agentes de conversação em aplicações multimédia. Pretende-se que os resultados sejam obtidos através do *feedback* proporcionado por utilizadores tanto através do interface da ferramenta, como por resposta a um questionário.

Os objetivos específicos deste trabalho são os seguintes:

- Identificar a tecnologia adequada e desenvolver uma plataforma servidor-cliente que permita a avaliação de agentes de conversação.
- Escolher e implementar um conjunto de funcionalidades que simulem comportamentos humanos através de elementos multimédia como a sintetização de voz e representação visual através do uso de um Avatar.
- Realizar testes com vários utilizadores, utilizando diferentes configurações da plataforma de avaliação e recolher dados através da interação do utilizador com a ferramenta de teste e um questionário realizado antes e depois do teste.
- Analisar os dados recolhidos, determinando o grau de relação entre a eficiência das várias funcionalidades e o perfil do utilizador e se existem estratégias que sejam importantes para o uso de agentes de conversação em aplicações multimédia.
- Propor diferentes funcionalidades que possam ser utilizadas em desenvolvimento posterior sobre a plataforma de avaliação e identificar possíveis adaptações de forma a permitir o uso de agentes capazes de desempenhar tarefas mais específicas.

## 1.4 Estrutura do Relatório

Este trabalho encontra-se estruturado em seis capítulos dos quais, o primeiro é composto por esta introdução ao trabalho.

No segundo capítulo, é efetuada uma breve descrição da importância do trabalho de Alan Turing e como este seria fundamental para o aparecimento dos Agentes de Conversação tal como os conhecemos hoje.

No terceiro capítulo é apresentado o conceito de Agente de Conversação utilizado em aplicações multimédia e discutido como tem vindo a ser analisada a sua eficiência. Neste capítulo, pretende-se identificar quais as características de um Agente, que podem ser melhoradas para permitir uma interação mais eficiente com o utilizador.

No quarto capítulo é descrita a implementação da aplicação desenvolvida no âmbito de este trabalho. É fundamentada a tecnologia escolhida e são analisadas algumas das questões que foram consideradas durante o desenvolvimento da aplicação. São também docu-

mentadas as diversas possibilidades de alteração e configuração da plataforma desenvolvida.

No quinto capítulo são analisadas as questões relacionadas com o levantamento de dados através de vários testes experimentais. Para além dos critérios utilizados para a elaboração das questões a apresentar aos utilizadores, são também discutidos os resultados dos inquéritos efetuados.

O último capítulo contém as conclusões gerais deste trabalho, analisa os seus principais resultados e apresenta algumas perspetivas de desenvolvimentos futuros.

## Capítulo 2

# 2. Agentes de Conversação

Um Agente de Conversação é um programa de computador criado com a finalidade de imitar a capacidade humana de comunicar. Ou seja, é um programa que dada uma frase tenta formular uma resposta adequada, tal como faria um humano. Apesar de estes se terem desenvolvidos inicialmente com o intuito de explorar as semelhanças entre as capacidades de um humano e uma máquina, foram entretanto, objeto de diversos estudos de forma a poderem ser usados em diversos outros tipos de aplicações.

Neste capítulo, pretende-se então, analisar o contexto do aparecimento dos Agentes de Conversação, a sua evolução para o uso em aplicações Multimédia e quais os caminhos que esta tecnologia poderá tomar de futuro.

## 2.1 O trabalho de Alan Turing

Alan Turing (1912-1954) foi um matemático notável, considerado um dos pais da Ciência de Computadores. A sua vida curta e fascinante foi retratada em diversas obras de arte e ficção. Enquanto criptógrafo foi responsável por decifrar as comunicações dos submarinos Alemães, permitindo uma vantagem significativa das Forças Aliadas nas batalhas navais durante a segunda guerra mundial.

Apesar de grande parte do seu trabalho assentar sobre questões fundamentais da computação, como o próprio conceito de computabilidade e as suas limitações, e nunca se definir como um filósofo, Turing seria responsável por uma das obras mais citadas na literatura de filosofia moderna. [*Stanford Encyclopedia of Philosophy*]

### 2.1.1 A máquina de Turing

Antes de se falar sobre Agentes de Conversação, torna-se necessário contextualizar o seu aparecimento. Um dos trabalhos mais notáveis de Alan Turing é, de facto, a sua tese, elaborada em conjunto com Alonzo Church, que define os limites da computabilidade relacionando-os com um modelo matemático descrito por si, a máquina de Turing.

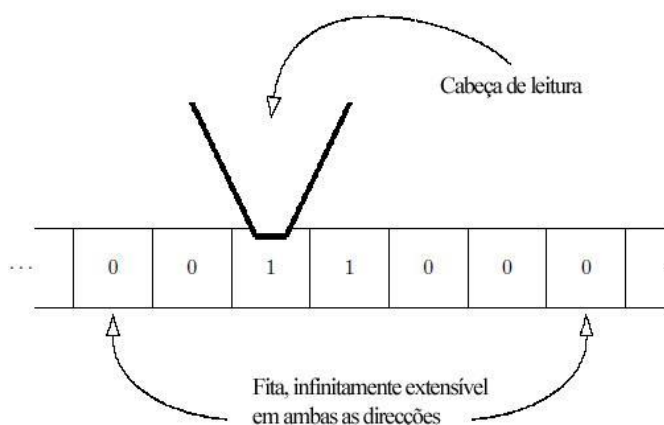


Esta máquina, não é mais do que um modelo matemático que representa uma fita de tamanho infinito e de um conjunto simples de operações. Estas operações são efetuadas por uma cabeça capaz de ler e escrever em uma das células da fita tal como exemplificado na Ilustração 1. [Stanford Encyclopedia of Philosophy]

De acordo com a tese de Church Turing, através de um número limitado de operações e da manipulação e armazenamento de símbolos, a máquina de Turing permite efetuar todos os conjuntos de operações que possam ser efetuadas através de um método definido, isto é, descritas por um algoritmo.

Turing seria então responsável por identificar e definir o conjunto de números computáveis e operações que podem ser efetuadas sobre estes.

Apesar de Turing não ter sido o primeiro a apresentar uma noção geral de computabilidade, terá sido o primeiro a publicar, em 1937, uma definição convincente que atualmente continua a ser verificada e utilizada. [Cooper 2004]



*Ilustração 1. Exemplo de uma fita infinita e cabeça de leitura tal como descrito por Turing [Cooper 2004]*

É possível encontrar em todo o trabalho de Turing uma tentativa de aproximação entre a computação feita por uma máquina e as capacidades humanas. De facto, a própria máquina de Turing foi desenhada considerando operações que poderiam ser desempenhadas por uma calculadora humana com um número finito de equipamento mas sem limite de papel e tempo disponível. As operações que podem ser efetuadas pela máquina podem ser relacionadas com ações humanas. É também possível estabelecer uma analogia entre o número finito de estados que a máquina de Turing pode tomar e a memória finita de uma calculadora humana. [Epstein, Roberts e Beber 2008]

Convém salientar que o trabalho de Turing antecede o aparecimento do computador tal como o conhecemos e o seu trabalho seria bastante importante no sentido em que definiu os fundamentos teóricos sobre os quais estes viriam a assentar. De facto, uma máquina de

Turing representa um conjunto de instruções, ou seja, o equivalente a um programa de computador. Desta forma, podemos estabelecer uma analogia entre a Máquina Universal de Turing que representa o conjunto de todas as instâncias de Máquinas de Turing e o computador atual. [Stanford Encyclopedia of Philosophy]

O fascínio de Turing em tentar relacionar a capacidade de máquinas capazes de executar diferentes operações de computação com diferentes capacidades humanas viria a tornar-se mais evidente através de uma ideia apresentada 15 anos mais tarde. [Epstein, Roberts e Beber 2008]

### 2.1.2 O jogo da Imitação

Em 1950 é descrito por Alan Turing um teste denominado de Jogo da Imitação. A finalidade deste teste é determinar se é possível simular o pensamento humano através do uso de uma máquina, ou seja, se uma máquina é capaz de pensar. Este teste é descrito como um jogo cuja finalidade é levantar a questão “Uma máquinas consegue pensar?”.

De acordo com Turing, a máquina será capaz de reproduzir a inteligência humana quando conseguir estabelecer um diálogo consistente sem ser distinguida do ser humano. Tal como proposto por Turing em “*Computing machinery and intelligence*”, a pergunta “As máquinas conseguem pensar?” seria substituída por outra de forma a torná-la mais objetiva.

O problema seria colocado na forma de jogo recorrendo a três participantes, um homem (A), uma mulher (B) e outra pessoa, de qualquer sexo, para desempenhar a tarefa de interrogador. O interrogador seria colocado em uma sala diferente das outras duas pessoas. O objetivo do jogo seria o interrogador identificar qual o participante masculino e o feminino. O interrogador poderá colocar perguntas a ambos participantes às quais irá receber a resposta correspondente.

Para que o tom de voz não tenha influência, estas perguntas devem ser feitas por escrito sem haver contacto entre os participantes e o interrogador. O objetivo do jogo ganha sentido quando se substitui o jogador A por uma máquina. De acordo com Turing, esta experiência poderia ser efetuada idealmente recorrendo ao uso de uma tele-impressora.

A questão fundamental, de acordo com Turing, é saber se o interrogador terá a mesma taxa de sucesso ao substituírmos o homem por uma máquina, substituindo assim a pergunta “as máquinas conseguem pensar?”. [A. M. Turing 1950]

A questão levantada por Turing, tem sido bastante debatida devido ao seu carácter multidisciplinar e foi um dos principais impulsionadores do início do estudo da inteligência artificial. Apesar de contestada, por ser considerada uma simplificação excessiva que não responde completamente à questão inicial, é necessário entender o valor do problema formulado por Turing pois tem a vantagem de simplificar e estabelecer uma distinção entre as capacidades físicas e intelectuais do homem como consequência de não existir contacto

físico ou visual entre o interrogador e os participantes. Ou seja, é necessário entender a questão colocado por Turing como um modo de simplificar e permitir a comparação entre a capacidade humana e a de uma máquina, restringido a questão à capacidade de comunicar. [Epstein, Roberts e Beber 2008]

Recorrendo à tecnologia atual, é possível um juiz estabelecer comunicação escrita com outra pessoa e com uma máquina sem conhecer a identidade da fonte da comunicação, através de terminais ligados em rede.

### 2.1.3 O Prémio Loebner

Em 1991, com o intuito de incentivar o interesse para a pesquisa em essa área, é iniciada por Hugh Loebner uma competição anual com a finalidade de premiar o melhor desempenho no teste inicialmente descrito por Turing [*Loebner Prize*]. Os concorrentes de esta competição são avaliados por dez júris, selecionados em resposta a um anúncio no jornal. Após comunicação escrita com duas entidades, uma humana e outra um agente de conversação, tal como descrito por Turing, cada elemento do júri atribui uma classificação de acordo com a sua convicção que conseguiu identificar o elemento não humano. O primeiro prémio desta competição será apenas atribuído quando um agente de conversação conseguir enganar um terço dos elementos do júri.

Apesar da decisão de cada elemento do júri corresponder a uma resposta de Verdadeiro ou Falso, é pedido aos júris para ordenar os concorrentes de acordo com o seu desempenho obtendo-se um formulário com os resultados semelhante ao exemplificado na Ilustração 2. Desta forma, é possível classificar os vários concorrentes de acordo com a sua posição média e atribuir o prémio de vencedor do torneio. [Shieber Stuart 1993]

Máquinas					Humanos				
1	2	3	4	5	6	7	8	Menos Humano	Mais Humano
<i>B</i>	<i>A</i>	<i>E</i>	<i>D</i>	<i>C</i>	<i>F</i>	<i>H</i>	<i>G</i>		

*Ilustração 2. Exemplo do formulário utilizado para classificar os concorrentes na primeira edição da competição Loebner. [Shieber Stuart 1993]*

#### 2.1.3.1 As primeiras edições do Prémio Loebner

Convém também referir que as regras da competição foram pensadas ao longo de dois anos. De forma a manter o interesse da competição, protegendo as máquinas de uma eliminação rápida, foram pensadas duas regras. Uma delas seria limitar o tema da conversa a um tópico escolhido pelo participante. Caso contrário, devido à quantidade de tópicos pos-

síveis para uma conversa, um júri seria capaz de descobrir qualquer concorrente não humano em um curto espaço de tempo.

No caso da primeira edição do prémio Loebner, os seis temas escolhidos para serem utilizados por agentes de conversação foram vinhos Borgonha, Martinis secos, conversas de circunstância, conversas caprichosas e insatisfação em relações.

Algumas das escolhas, apesar de adequadas a diversos contextos sociais, não definem propriamente um tema de conversa mas sim um estilo de conversação com o intuito de justificar respostas demasiado simples ou mesmo inadequadas. Talvez, por isso mesmo, a escolha de um tema como conversas caprichosas se tenha revelado fundamental para predefinir o agente de conversação com melhor desempenho, pois justifica como caprichosas e impulsivas grande parte das respostas que podem ser aleatórias ou mesmo desconexas, dificultando a tarefa do júri.

No caso dos dois participantes humanos, os tópicos de conversação escolhidos foram Shakespeare e Vestuário Feminino.

Outra regra importante restringe a possibilidade de os júris recorrerem a comunicação inadequada para tentar obter uma resposta do participante. Ou seja, os júris são obrigados a escrever frases com sentido relacionadas com o tópico e nunca tentar obter uma resposta enganando o agente através de frases sem sentido. [Epstein Robert et al 2008] [Shieber Stuart 1993]

Apesar de a utilização de regras para proteger o desempenho das máquinas, rapidamente se verificou a facilidade com que alguém com conhecimento sobre o funcionamento destes programas consegue identificar facilmente o agente, pois bastantes vezes as técnicas utilizadas pelos agentes tornam-se bastante evidentes por estarem descontextualizadas.

É possível verificar estas ocorrências em excertos retirados da primeira edição da competição.

*“01:11:42 Júri 10*

*Tem sido um longo dia! Agora penso que estou pronto para alguma diversão. O tempo não está muito agradável hoje. Tens algumas anedotas que possam levantar as nuvens?*

*01:13:01 Terminal 5*

*Toda a gente fala sobre o tempo mas ninguém parece fazer muito acerca disso. Porque tens de me dizer agora que pensas estar pronto para alguma diversão? Suspiroooooo.....”* [Shieber Stuart 1993 retirado de *Cambridge Center for Behavioral Studies* 1991 tradução nossa]

Na primeira edição, o Agente com melhor desempenho conseguiu ser classificado como humano por dois elementos do júri. Surpreendentemente, um dos participantes humanos

foi identificado duas vezes, de forma errada, como um computador devido ao seu desempenho notável a responder sobre questões de Shakespeare.

De acordo com a previsão de Turing, a partir do ano 2000 um computador seria capaz de ultrapassar o seu desafio com alguma solidez e conseguiria enganar 70% dos júris durante 5 minutos. Atualmente ainda não se verificou a previsão feita por Turing e o objetivo de conseguir ultrapassar o teste, sem restrições, parece ainda bastante longínquo. [Shieber Stuart 1993]

Considerando as limitações estabelecidas pelas regras podemos afirmar que de facto se perde a essência da ideia de Turing. A ideia de avaliar se uma máquina é capaz de comunicar inteligentemente implica que esta não deve estar limitada por regras que atuem sobre a própria comunicação.

### **2.1.3.2 O Prémio de Loebner sem restrições**

A partir de 1994, deixou de haver restrições na competição Loebner de forma a permitir uma avaliação dos agentes de conversação adequada à proposta inicial de Turing. Desta forma passa a ser possível aos júris utilizarem técnicas mais sofisticadas para procederem à identificação dos programas concorrentes. Da mesma forma que as técnicas de deteção por parte de alguns elementos do júri se tornaram mais elaboradas, os agentes de conversação também começaram a ser implementados com novas técnicas para conseguir enganar os elementos do júri.

Algumas questões sobre a validade do formato de Prémio Loebner, mesmo sem restrições, viriam a ser levantadas por vários motivos. Para além de não existirem progressos significativos na tecnologia de Agentes de Conversação, a qualidade dos agentes a participar no Prémio Loebner também se viria a revelar em decréscimo. A motivação de um prémio monetário funciona também como um incentivo a todo o tipo de técnicas elaboradas para explorar as regras da competição e enganar os júris. [Shieber Stuart 1993]

Em 1996, foi explicado pelo participante vencedor as técnicas que havia utilizado para vencer a competição. Este agente, para além de explorar a possibilidade de interagir o menos possível com o júri durante os 30 minutos da prova, recorria também a técnicas bastante ingénuas e nada relacionadas com a tecnologia em questão. Através de um início de conversação devidamente planeado e implementado sem recorrer a técnicas de Inteligência Artificial, este agente conseguia conduzir a conversa de forma a obter os tipos de conversa onde tem mais hipótese de sucesso.

Uma das principais motivações para a continuidade do Prémio Loebner, surge em 2008 com o agente vencedor a conseguir enganar três dos doze elementos do júri. No caso de este agente ter conseguido enganar mais um elemento, um quarto do total, seria o primeiro participante a receber a medalha de ouro.

É também de salientar que atualmente é possível verificar uma diminuição da documentação sobre o Prémio Loebner e sobre as várias formas de abordar o problema. Possivelmen-

te este facto explica-se pela diminuição do interesse em esta competição e também de forma a garantir que a tecnologia, associada Agentes mais competitivos, possa ser utilizada com fins comerciais.

Convém salientar, que existe a possibilidade de o torneio elaborado por Loebner não ser o mais adequado para dinamizar os esforços na questão colocada por Turing. De facto, apesar de esta ser importante para colocar alguma atenção sobre o assunto, pode-se dizer que o esforço de os participantes se centra em desenvolver um agente capaz de ganhar a competição.

## 2.2 Agentes de Conversação

Apesar de existirem diversas implementações de Agentes de Conversação, grande parte do desenvolvimento foi efetuado em torno das ideias utilizadas na criação do primeiro agente de conversação, o *ELIZA*.

### 2.2.1 ELIZA, o Psicoterapeuta Rogeriano

Desde que foi proposto o teste de Turing, uma das primeiras tentativas de ultrapassar este desafio foi desenvolvida por Joseph Weizenbaum, um programa denominado de *ELIZA* [Weizenbaum 1966]. Este programa continua a servir de base para muitos dos agentes de conversação desenvolvidos ainda hoje. Alguns dos vencedores do Prémio Loebner são variantes do *ELIZA* [Loebner Prize].

De acordo com Joseph Weizenbaum, o funcionamento do programa é bastante simples e baseia-se essencialmente na deteção de uma palavra-chave que irá determinar o conjunto de alterações aplicados à frase de inicial. Depois de ser calculada a resposta é então exibida. Ou seja, o Eliza não aborda a questão de simular a inteligência humana através de técnicas elaboradas de inteligência artificial mas sim recorrendo a um conjunto simples de regras e manipulação de texto, criando uma ilusão de inteligência.

Existem outros pormenores técnicos relacionados com este mecanismo de transformação de dados para que possa funcionar de forma eficiente. Apesar de o funcionamento do programa ser pouco mais que uma questão de manipulação de texto devem ser consideradas as seguintes questões enumeradas por Joseph Weizenbaum em “*Eliza: a computer program for the study of natural language communication between man and machine*”

1. Identificação da palavra mais importante na mensagem de entrada em conjunto com um contexto na qual esta possa aparecer.
2. Escolha e execução da transformação apropriada tal como definido pelas regras associadas à palavra-chave escolhida.
3. Criação de um mecanismo que permita uma resposta inteligente quando o texto de entrada não contém palavras-chave.

4. Provisão de um mecanismo que permita a edição e extensão das transformações que definem as respostas.

Uma das vantagens desta implementação é a vantagem de não estar associada exclusivamente a nenhuma língua. Assim sendo, pode ser desenvolvido um conjunto de padrões e transformações, isto é, um *script* adequado a qualquer língua, o que permitiu o aparecimento de implementações tanto em Galês como Alemão ou Inglês.

O *script* mais utilizado do *ELIZA* caracterizava-se por agir como um psicoterapeuta Roge-riano, isto é, que tenta centrar a conversação sobre a perceção do utilizador dos diversos temas. Esta abordagem permite exibir a falta de conhecimento sobre assuntos bastante comuns de forma a parecer intencional, respondendo com uma pergunta.

*Como está? Por favor diga o seu problema.*

*>> Passo demasiado tempo com computadores.*

*Porque menciona computadores?*

*>> Tu és um.*

*Fica satisfeito em acreditar que sou um?*

*>> Sim.*

*Parece estar bastante certo.*

*>> Suponho que estou.*

*Há quanto tempo tem estado?*

[Güzeldere e Franchi 1995 tradução nossa]

No exemplo é possível ver a forma como o agente utiliza perguntas para manter a simplicidade da conversa. É também possível observar uma das limitações da estratégia utilizada por este tipo de agente de conversação quando este utiliza uma base de conhecimento que não antevê certas ocorrências. Neste caso, o Agente esperava encontrar uma palavra a seguir ao verbo sou pelo que não foi capaz de responder coerentemente.

No Anexo 1 é possível encontrar excertos de uma conversação entre o Agente *ELIZA* e uma paciente.

### 2.2.2 A.L.I.C.E.

Uma destas variantes é o A.L.I.C.E. ou *Alicebot* que também se baseia em mecanismos de detecção de padrões em conjunto com uma base de conhecimento em *AIML* para elaborar as suas respostas [Alice]. Apesar de este ser bastante semelhante ao seu antecessor, o uso de ficheiros estruturados de acordo com a norma proposta por Dr. Richard S. Wallace permitiu o aparecimento de várias variantes deste agente de conversação e o desenvolvimento de bases de dados de conhecimento em várias línguas e sobre diversos temas específicos.

A estrutura do Agente de Conversação *ALICE* consiste em 41000 elementos denominados de categorias. A cada uma destas categorias encontra-se associado um padrão que repre-

senta uma pergunta e a respetiva resposta. Esta informação é guardada em *AIML* ou seja, em uma estrutura em árvore. A correspondência entre o texto de entrada e os padrões definidos no *AIML* são geridos por um objeto bastante compacto em termos de utilização de memória e recorrendo a um algoritmo eficiente que permite tempos de procura bastante rápidos.

O objeto responsável pela organização de toda a informação referente às categorias é denominado de *Graphmaster*. Esta informação é guardada em uma estrutura em árvore, o que permite evitar a duplicação de informação conseguindo assim uma compressão eficiente da informação.

```
<aiml>
  <category>
    <pattern> OLA </pattern>
    <template> Ola! Como estás? </template>
  </category>
</aiml>
```

No caso do exemplo anterior, este utiliza a frase “OLA” para identificar uma saudação e responde de forma adequada. Este é um dos exemplos mais simples das possibilidades do *AIML* e tem a desvantagem de responder exatamente a apenas uma frase.

Através de utilização de um conjunto de regras recursivas e do operador “\*”, é possível utilizar diversas técnicas que permitem um uso mais versátil do *AIML*.

Uma das técnicas é a redução simbólica que permite a simplificação de formas gramaticais complexas para poderem ser tratadas pelas regras mais simples. A tag <srai> é responsável pela recursão da regra definida.

```
<aiml>
  <category>
    <pattern> OLA TUDO BEM? </pattern>
    <template><srai>OLA</srai></template>
  </category>
</aiml>
```

No caso do exemplo dado, também poderíamos ter definido uma regra onde as várias partes da frase fossem tratadas através de regras diferentes. Ou seja, em vez de reduzirmos, limitamo-nos a dividir a frase e ambas as partes são tratadas recursivamente.

```
<aiml>
```



```

<category>
  <pattern> OLA * </pattern>
  <template><srai>OLA</srai><srai><star/></srai></template>
</category>
</aiml>

```

Em este exemplo, a palavra “OLA” iria ser tratada pela regra definida anteriormente e o resto da frase iria tentar encontrar uma regra adequada. Considerando que não existe uma regra que identifique a frase que pretendemos, é possível escolher uma resposta para o conjunto de frases ou variáveis não definidas.

```

<unknown>Não compreendi. Podes explicar melhor?</unknown>

```

Um dos aspetos mais importantes na escrita de *AIML* eficiente são as categorias que funcionam por defeito. Estas são as categorias que incluem o carácter ‘\*’ e que uso de outras categorias através da *tag* <srai>.

Uma parte significativa das interações com utilizadores irá de facto ser corresponder a estas categorias devido à impossibilidade de escrever regras específicas para todas as possibilidades. De facto, uma regra com bastante importância é a última das categorias por defeito que é definida por apenas pelo operador representado pelo carácter ‘\*’ operador vulgarmente usualmente de *wildcard*.

Normalmente esta categoria é composta por um conjunto longo de frases aleatórias que permitam redirecionar a conversa para assuntos que o Agente de Conversação tenha mais conhecimento.

```

<category>
  <pattern>*</pattern>
  <template><random>
    <li>Quantos anos tens?</li>
    <li>Qual é o teu signo?</li>
  </template>
</category>

```

Podem ser implementadas mais categorias por defeito, utilizando palavras e o operador *wildcard*.

```

<category>
  <pattern>Preciso de ajuda *</pattern>

```

*<template>Pergunta o que quiseses que eu ajudo</template>*

*</category>*

Desta forma é possível responder a várias opções como “preciso de ajuda com o Word” ou “preciso de ajuda para mudar de casa”. Com a resposta dada, o agente consegue exibir um comportamento humano demonstrando preocupação e disponibilidade e ao mesmo tempo tenta forçar o utilizador a utilizar questões diretas que possam estar definidas em categorias específicas.

Tal como podemos verificar, a criação de respostas eficientes em AIML, exige mais criatividade e a capacidade de antever comportamentos do que conhecimentos técnicos.

Uma das questões associadas ao facto de uma máquina tentar imitar um humano, é a quantidade de informação necessária para esta poder responder a questões sobre os mais diversos temas. Com o uso da estrutura utilizada pelo agente *ALICE*, é possível utilizar um conjunto de bases de dados de conhecimento desenvolvidas paralelamente. Outra possibilidade é a criação de ficheiros *AIML* de forma automática, recorrendo a técnicas de *targeting*. [Weizenbaum 1996]

A possibilidade de utilizar o *targeting*, isto é, permitir o aumento da base de dados de conhecimento do agente, é de facto, um dos pontos fortes desta implementação. Esta técnica consiste na utilização normal do Agente de Conversação, e no caso de se obter uma resposta insatisfatória, a pessoa responsável pela administração do Agente, o *botmaster*, pode identificar qual seria a resposta adequada naquela circunstância. Desta forma, o desempenho do Agente encontra-se associado à capacidade de gestão eficiente do conhecimento do agente por parte do *botmaster*.

*“Utilizador: Onde encontro uma página web com Australopitecos?”*

*Alice: Onde a queres encontrar?*

*Utilizador: Má resposta Alice, hora de Aprender uma Linha nova*

*Alice: Como responderias a ‘Onde encontro uma página web com Australopitecos?’*

*Utilizador: Eu diria: ‘<http://www.ancientskulls.net/australopithecus.html>’*

*Alice: Obrigado, humano. Tenta perguntar-me outra vez*

*Utilizador: Onde encontro uma página web com Australopitecos?*

*Alice: <http://www.ancientskulls.net/australopithecus.html>”*

*[The Elements of AIML Style tradução nossa]*

### 2.2.3 PARRY, o paciente Paranoico

Uma das implementações de agentes de conversação que surgiu pouco depois do *ELIZA* foi o *PARRY*. Este programa caracteriza-se por tentar simular um discurso com características de um comportamento paranoico. Isto é, ao contrário do *ELIZA*, tenta mostrar uma personalidade.

O *PARRY* surge como crítica à falta de capacidade do *ELIZA* de utilizar um modelo que represente o seu estado interno. Assim, o *PARRY* elabora as suas respostas de acordo com o seu estado interno, isto é, as suas respostas dependem do seu estado emocional. Apesar de ter sido criticado pela simplicidade da abordagem, convém salientar que este método não é trivial. [Colby 1975]

*PARRY* foi capaz de enganar um conjunto de psicanalistas, que comunicaram com ele através de um terminal de escrita sem conhecerem a sua identidade. Isto é, Kenneth Mark Colby conseguiu elaborar uma história em que reproduziu com eficácia o Jogo da Imitação de Turing, e obteve por parte de pessoas habilitadas, a confirmação que estavam a comunicar com um paciente paranoico. A ideia de *PARRY* ter sido o primeiro agente a ter sucesso no Teste de Turing, foi prontamente criticada.

Para fundamentar a sua prova, Colby, demonstrou a eficiência da utilização do modelo emocional do seu agente de conversação, contrastando com a diferença de um agente que utiliza as mesmas respostas mas de forma aleatória, isto é sem recorrer ao seu modelo interno para decidir qual a escolha mais adequada. [Mauldin 1994]

É também interessante observar, por mera curiosidade, um excerto resultante de diálogos entre *ELIZA*, o psicoterapeuta e *PARRY*, o paciente paranoico.

*PARRY: Eu sei que a máfia controla as redes de corrupção*

*ELIZA: Sobre quem estás mesmo a falar?*

*PARRY: Alguma vez pensaste realmente sobre a Máfia?*

*ELIZA: Supõe que alguma vez pensaste realmente sobre a Máfia?*

*PARRY: Vamos falar sobre outra coisa.*

*ELIZA: Não tenho a certeza que te entendo completamente*

*PARRY: Continuas a perguntar as mesmas coisas, porque não trocar de tópico?*

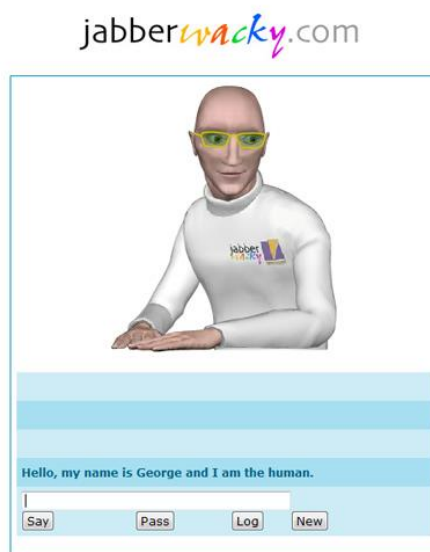
*ELIZA: Em que sentido?"* [Güzeldere e Franchi 1995 tradução nossa]

### 2.2.4 Jaberwacky

Outro agente de conversação digno de se mencionar, devido às suas prestações vencedoras em edições mais recentes do Prémio Loebner, é o *Jabberwacky*. Apesar de este projeto ter sido iniciado em 1981, foi nas edições de 2005 e 2006 do Prémio Loebner que este reve-

lou a eficiência dos seus métodos. Esta implementação distingue-se pelo facto de utilizar diversas técnicas que permitem a aprendizagem do agente sendo assim possível a obtenção de uma resposta influenciada pelo conhecimento obtido tanto na conversa atual como conversas anteriores.

Este agente de conversação, encontra-se implementado, em conjunto com uma aparência virtual e sintetização de voz, com o nome de *George* e, tal como os agentes referidos anteriormente, encontra-se disponível na Internet para experimentação com um interface semelhante à Ilustração 3.



*Ilustração 3. Interface do Agente de Conversação George*

Contudo, convém referir que a tecnologia associada a este agente é propriedade comercial e como tal, os detalhes do seu funcionamento não se encontram disponíveis para estudo. Devido à facilidade com que este agente é capaz de enganar utilizadores humanos, o seu criador prevê que esta tecnologia possa a ser vir largamente implementada e bastante prática se associada a repositórios de informação bastante desenvolvidos.

*“jabberwacky: És um extraterrestre?”*

*user: Sim um extraterrestre cor-de-rosa.*

*jabberwacky: Não exactamente.*

*user: Sim, sou um extraterrestre e sou cor-de-rosa.*

*jabberwacky: És humano?*

*user: Metade. Sou metade humano e metade extraterrestre.*

*jabberwacky: Isso é estranho.*

*user: Sim. Eu gosto de ser estranho.*

*jabberwacky: Bem, és bastante bom nisso.*

*user: Obrigado.*

*jabberwacky: Não é um elogio.*

*user: Mas eu levo como um elogio.*

*jabberwacky: Muito bem.*

*user: Obrigado.”*

[<http://jabberwacky.icogno.com/j2convbycategory> tradução nossa]

## 2.2.5 Elbot

Outro agente digno de destaque é o *Elbot*. Desenvolvido por Fred Roberts utilizando tecnologia da empresa Artificial Solutions [*Artificial Solutions*]. Este agente seria o vencedor da edição de 2008 do Prémio Loebner [*Loebner Prize*]. Tal como referido anteriormente, este foi o Agente de conversação que teve mais perto de obter a medalha de Ouro no Prémio Loebner. A sua boa prestação no Prémio Loebner foi conseguida utilizando tecnologia de Interação com Linguagem Natural denominada de *Teneo*. [Elbot the Robot]

Como este agente foi desenvolvido utilizando tecnologia comercial, não podem ser estudados os pormenores relativos ao seu funcionamento, contudo convém salientar a importância de abordagens que não assentam exclusivamente sobre tecnologia desenvolvida anteriormente podendo assim ser importantes no aparecimento de novas técnicas.

Outro facto importante relaciona-se com a quantidade de produtos comerciais disponibilizados recorrendo a esta tecnologia o que pode servir com uma motivação adicional para a pesquisa e desenvolvimento nesta área.

## 2.2.6 Suzette

O Agente de Conversação Suzette merece uma breve referência por ter sido o último Agente de Conversação a ganhar o Prémio Loebner até à data de elaboração deste documento. Este Agente representa assim, o estado atual da tecnologia de Agentes de Conversação apesar de ter obtido uma classificação inferior à do Agente vencedor da edição anterior.

O Agente de Conversação *Suzette* foi vencedor da edição de 2010 conseguindo enganar 1 dos 4 elementos do Júri. [*Loebner Prize*] Apesar de se verificar um retrocesso em relação ao desempenho do agente vencedor na edição anterior, convém realçar o facto de a tecnologia utilizada também tentar utilizar uma alternativa ao AIML.

A abordagem deste agente tenta contornar o facto de as implementações que utilizam AIML necessitarem de muita informação para fazerem muito pouco. Outra limitação do AIML é que quando utilizamos uma base de conhecimento muito extensa, a sua organização torna-se bastante difícil. A tecnologia do Agente *Suzette*, que se centra sobre uma linguagem de *scripting* denominada de CHAT-L, tenta reduzir a redundância de informação e antevê mecanismos de visualização que permitam uma representação mais compreensível dos dados. [*Beyond AIML 2008*]

Um das características que distingue o CHAT-L das outras implementações é a forma como este interpreta a informação recebida. Tal como já foi referido, o *AIML* descarta toda a

pontuação ignora também as capitalizações, o que resulta em perda informação enquanto o CHAT-L tenta utilizar esta informação para simplificar a base de conhecimento.

Outra das características do CHAT-L é introduzir o conceito de tópico que é fundamental no seu funcionamento. De facto, todas as regras têm de estar associadas a tópicos, de acordo com Bruce Wilcox, um tópico como cinema tem cerca de 200 regras associadas. É também possível definir uma lista de palavras associadas a cada tópico. Este conceito permite que o Agente determine qual o tópico que está a ser discutido o que permite uma redução do número de regras a verificar. Como é guardado o historial dos tópicos discutidos, é também possível usar esta informação para retomar tópicos anteriores em vez de utilizar respostas repetidas quando estas se esgotam. [Wilcox 2011]

## 2.3 Conclusão

Apesar de ainda não existir tecnologia de agentes de conversação capaz de simular com eficiência a capacidade de conversação humana tal como idealizado por Alan Turing, é possível verificar uma melhoria significativa em esta tecnologia tal como se tem verificado com os resultados do prémio Loebner. [*Loebner Prize*]

Verifica-se também a importância das condições utilizadas na avaliação de Agentes de conversação, isto é, a importância de restringir as conversações a um tema específico. [Shieber Stuart 1993]

Outra conclusão importante é que, apesar de persistir a utilização de tecnologia baseada no Agente Alice e a sua base de conhecimento *AIML*, mais recentemente tem-se vindo a verificar o sucesso de Agentes que utilizam diferentes abordagens e tecnologia. [*Loebner Prize*]

## Capítulo 3

# 3. Agentes de Conversação em Ambientes Virtuais

Com o desenvolvimento da tecnologia e o aumento de implementações disponíveis, os agentes de conversação foram evoluindo para além da questão levantada por Turing. Isto é, atualmente é possível aumentar a credibilidade e as funcionalidades do agente de conversação, através do uso de multimédia.

É possível encontrar Agentes de Conversação a desempenhar diversas funções. Para além do uso como entretenimento, podem ser utilizados para fornecer apoio a clientes de diversos serviços comerciais. Outra possibilidade que tem sido objeto de diversos estudos académicos é a de utilizar os Agentes de Conversação no auxílio ao ensino de diversas áreas de conhecimento. O uso de Agentes de Conversação em conjunto com o conhecimento atual sobre psicologia educacional tem a possibilidade de aumentar o interesse e a capacidade de retenção do educando [Adwin, Winne, Nesbit 2005].

De forma a tornar a experiência mais realista e agradável, é comum atribuir-se ao Agente de Conversação uma representação humana. Estes Agentes são normalmente conhecidos como Agentes de Conversação Personificados (ECA) e existem diversos estudos que pretendem relacionar a sua eficiência com as diversas características humanas que lhe podem ser atribuídas. [Beun, Vos, Witteman 2003]

Convém também salientar que para além do uso de informação associada às várias capacidades sensoriais humanas, fatores mais complexos como a emoção e a personalidade podem também contribuir para o aumento da credibilidade do agente. [Thalmann, 2005]

## 3.1 A influência da linguagem e do domínio de conhecimento

Existem diversas tentativas de utilizar Agentes de Conversação, mais especificamente, Agentes de Conversação Personificados com o intuito de melhorar a aprendizagem em diversas áreas. Tal como referido anteriormente, a possibilidade da utilização de bases de conhecimento sobre temas específicos em agentes como *ALICE* permite que este possa ser

adaptado a diversos contextos de interação e ensino. Desta forma, a capacidade de resposta do agente é determinada pela gestão de dados feita pelo gestor do agente, ou seja, o botmaster. Para além da quantidade e qualidade da informação guardada na base de conhecimento, a eficiência do agente é também determinada pelo próprio tema que pretendemos lecionar.

Para além do conhecimento técnico que o agente deve ser capaz de transmitir, isto é, o conhecimento sobre o domínio, é também importante a adição de conhecimento conversacional adequado ao contexto. Desta forma, o agente poderá, por exemplo, pedir a reformulação das perguntas de forma a serem adequadas ao que ele pode responder. Estes dados foram comprovados em estudos com um Agente elaborado para ensinar Geometria Euclidiana [Han e Kim, 2001] e referidos em estudos com um Agente de Conversação capaz de disponibilizar informação sobre telecomunicações. [Schumaker e Chen, 2006]

Existe também uma relação entre o tipo e o tamanho das perguntas e o índice de satisfação do utilizador com a resposta formulada pelo agente. [Schumaker e Chen, 2010]

### 3.1.1 AZ-ALICE e TARA

#### 3.1.1.1 AZ-ALICE

De forma a possibilitar o aumento da base do conhecimento do agente Alice de uma forma mais prática, foi idealizado o agente *AZ-Alice*. Este agente foi utilizado para testar a influência do aumento da base de conhecimento de um tema específico. Neste caso o agente pretendia ensinar e avaliar a retenção de informação relacionada com telecomunicações.

Para além dos componentes que são comuns ao agente Alice, o agente desenvolvido neste estudo, o *AZ-Alice* é dotado de um interface que permite a autenticação dos utilizadores. Desta forma, é possível efetuar comunicação personalizada pois o agente passa a ser capaz de reconhecer o nome do utilizador.

A diferença mais importante é o facto de o sistema *AZ-Alice* gravar toda a informação relacionada com as conversas e recolher através do interface a avaliação e alternativas propostas pelo utilizador para melhorar a resposta do agente. Este módulo de avaliação permite que a base de conhecimento seja melhorada conforme a utilização do sistema por utilizadores especializados no domínio em questão.

	Experiência 1	Experiência 2
<b>BaseBot</b>	74	98
<b>TeleComm</b>	91	113

Tabela 1. Número de participantes [Schumaker, Ginsburg, Chen e Liu 2006]

No estudo foi utilizado um agente *ALICE* com o *AIML* de base que é constituído por 24835 categorias de conhecimento, o agente *TeleComm* foram acrescentadas, a uma base de conhecimento idêntica, 298 definições acerca de telecomunicações.



Recorrendo a duas fases de experimentação, este estudo permitiu verificar a utilização de um agente de conversação como uma ferramenta eficiente para recolher conhecimento relacionado tanto como a forma como é efetuada uma conversação como com informações adicionais a um domínio específico. De acordo com os autores, a diminuição da taxa de correção, entre os vários estudos, prova que algumas fases adicionais de correção seriam o suficiente para elaborar uma base de conhecimento com capacidade de responder a perguntas relacionadas com o domínio das telecomunicações. Verifica-se também que a restrição a um domínio específico de conversação permite obter um grau de satisfação superior mesmo quando o agente não se encontra especialmente preparado para esse domínio.

Os autores de este estudo concluíram também que, no caso de ser permitido a todos os utilizadores propor uma correção para respostas incompletas do agente, convém antever que algumas sugestões serão erradas, tanto de forma involuntária como propositadamente. A desvantagem do aparecimento destes erros é que, apesar de serem facilmente eliminados, serão necessárias fases de correção adicionais. [Schumaker, Ginsburg, Chen e Liu 2006]

De forma aprofundar a análise sobre a utilização de agentes de conversação em domínios específicos, foram também feitos testes recorrendo a diferentes configurações do agente TeleComm. A utilização de um agente com uma base de conhecimento de conversação reduzida de 23735 para 3892 entradas, permite então comparar a eficiência de agentes com uma base de conhecimento de domínio completa mas com diferentes graus de conhecimento de conversação. No caso de ser utilizada uma base de conhecimentos sobre telecomunicações mais completa, poderia ter sido ainda mais reduzida a base de dados de conversação. As 3,892 entradas de *AIML* escolhidas do conjunto base desenvolvido por Richard Wallace pretendiam responder a 65% das perguntas.

	<b>AIML Standard</b>	<b>AIML Telecomunicações</b>	<b>Total</b>
<b>BaseBot</b>	23735	0	23735
<b>TeleComm (Completo)</b>	23735	298	24032
<b>TeleComm (Limitado)</b>	3892	298	4190

Tabela 2. Distribuição das categorias AIML [Schumaker, Ginsburg, Chen e Liu 2007]

Em fase experimental, todos os participantes foram convidados a interagir apenas com uma das três variantes disponíveis, como diferentes partes de um curso introdutório de gestão de sistemas de informação. No fim de uma sessão de meia hora, os participantes teriam então de responder a um inquérito de forma a quantificar a satisfação de cada resposta dada pelo sistema. Foi também dada uma oportunidade ao aluno de propor melhorias às respostas obtidas. Apesar de não ser obrigatório, foi pedido aos alunos para manterem a conversação relacionada com o tema das Telecomunicações.

<b>Experiência 1</b>	
<b>BaseBot</b>	74
<b>TeleComm</b>	91
<b>Telecomm (limitado)</b>	92

Tabela 3. Número de participantes [Schumaker, Ginsburg, Chen e Liu 2007]

De acordo com os resultados obtidos experimentalmente, verifica-se que a utilização da componente conhecimento sobre o domínio é mais eficiente para obter conhecimento do que a utilização de apenas conhecimento conversacional. Isto deve-se à incapacidade de o conhecimento conversacional poder abranger todos os domínios apesar de cada domínio ter um número finito de definições e conceitos.

Outra conclusão obtida de este estudo é que, apesar do conhecimento conversacional não ser tão importante como o conhecimento de domínio, a utilização dos dois em conjuntos permite obter satisfação superior ao uso de conhecimento de domínio em exclusivo. Este resultado evidencia a importância de ambos os componentes e a necessidade de serem desenvolvidos de forma adequada para maximizar a eficiência de um agente de conversação quando utilizado em um domínio específico como as telecomunicações. [Schumaker, Ginsburg, Chen e Liu 2007]

### 3.1.1.2 TARA

Em um estudo semelhante ao do *AZ-Alice*, foi desenvolvido um agente capaz de responder a perguntas de ajuda em caso de um eventual ataque terrorista. Este agente, nomeado de *TARA*, possui diferenças em relação ao *AZ-Alice* no funcionamento sua base de conhecimento devido à forma como foram recolhidos os dados.

Foram acrescentadas à base de conhecimento do agente mais de dez mil categorias relacionadas com o tema. Esta informação foi recolhida em diversas páginas da Internet cujo conteúdo foi considerado adequado ao rigor exigido pela situação de uma emergência grave.

	<b>AIML Standard</b>	<b>AIML Terrorismo</b>	<b>Total</b>
<b>Agente base</b>	41873	10491	52354
<b>Agente de Conversação</b>	41873	0	41873
<b>Agente de Domínio</b>	0	10491	4190

Tabela 3 Distribuição das categorias AIML [Schumaker e Chen 2006]

Para testar a eficiência deste agente, foram efetuados testes recorrendo a três versões da base de conhecimento. Uma das versões correspondia à base de conhecimento que é utilizada normalmente com o agente *ALICE*, outra apenas com a informação recolhida e, finalmente, um agente que utiliza a totalidade do conhecimento. Desta forma seria possível determinar a eficiência do uso de informação relacionada com terrorismo quando utilizada em exclusivo e quando utilizada em conjunto com conhecimento que permita manter uma conversação.

	AIML Dialogo	AIML Domínio	Ambos
<b>Número de Utilizadores</b>	27	25	29
<b>Conforto</b>	22.2%	44%	31%
<b>Usabilidade</b>	11.1%	28%	31%
<b>Potencial Recomendação</b>	22.2%	28%	24.1%

*Tabela 4 Resultados obtidos com inquéritos de satisfação [Schumaker e Chen 2006]*

De acordo com a opinião dos participantes de este estudo, apesar de estes gostarem da amplitude de conhecimento e velocidade do sistema, mostram algum desagrado por receberem respostas em forma de definição. Ou seja, os utilizadores preferem uma interação com o sistema que se assemelhe mais ao fluxo normal de uma conversa. Este dado sugere a necessidade de haver uma filtragem e adaptação da informação recolhida de forma a ser apresentada ao utilizador de uma forma mais adequada a uma conversação.

Outra conclusão de este estudo é que os componentes de diálogo e de domínio funcionam melhor utilizados em conjunto do que quando utilizados de forma isolada.

Convém salientar que em ambos os estudos não foi possível verificar experimentalmente que exista uma clara vantagem em utilizar ambos os conjuntos de informação em simultâneo. Apesar de as respostas deixada pelos utilizadores em um questionário mencionarem os benefícios da utilização da versão mais completa do agente, este estudo deparou-se com o problema de possivelmente a informação relativa ao diálogo ter de ser mais cuidada para se adequar ao resto da informação. Este estudo debruçou-se também sobre a identificação de características comuns às perguntas que necessitam da utilização do conhecimento sobre o domínio em questão. [Schumaker e Chen 2006].

Ambos os estudos referidos anteriormente concluem essencialmente sobre a recolha de informação acerca de um tema específico e como esta deve ser apresentada ao utilizador.

Como tal, é possível utilizar os resultados dos dois estudos em conjunto para aprofundar sobre os vários tipos de questões que podem ser apresentados ao utilizador e quais as áreas de conhecimento que precisam de ser mais desenvolvidas para obter as respostas mais satisfatórias.

	Dialogo	Domínio	Total
<b>AZ-Alice - Dialogo</b>	23735	0	23735
<b>AZ-Alice - Domínio</b>	3892	298	4190
<b>AZ-Alice - Ambos</b>	23735	298	
<b>TARA - Dialogo</b>	41873	0	41873
<b>TARA - Domínio</b>	0	10491	10491
<b>TARA - Ambos</b>	41873	10491	52354

*Tabela 5 Distribuição das categorias AIML [Schumaker e Chen, 2010]*

Em ambos os estudos, cada utilizador interagiu em sessões de meia hora com duas variantes de apenas um agente e não houve nenhum utilizador que participasse nos dois estudos.

Apesar de ser utilizado em ambos os estudos a escala de *Likert* para identificar a satisfação do utilizador, o método de avaliação é feito de forma distinta. No caso do *AZ-Alice* a avaliação é feita no fim da conversação enquanto no *TARA* a avaliação é feita entre cada interação.

O tamanho da consulta parece, de facto, estar relacionado com a satisfação da resposta obtida. As consultas mais pequenas obtêm as respostas mais satisfatórias. Este resultado evidencia que deve haver preocupação em desenvolver mecanismos que permitam equilibrar a situação. Para melhorar a eficiência das respostas a consultas mais extensas, seria necessário incentivar o utilizador a utilizar frases curtas e simples ou simplesmente desenvolver uma base de conhecimento mais completa tendo em conta este tipo de consultas.

Este estudo identifica também algumas das ocorrências mais frequentes. No caso de este estudo, que utilizou a língua inglesa para o levantamento de resultados experimentais, foi possível verificar que em este tipo de agentes de conversação, grande parte das questões que lhe são feitas estão associadas a palavras iniciadas por “wh”, ou seja, o padrão “wh\*”. A ocorrência mais comum nas questões colocadas ao agente, em Inglês, é de facto a palavra “*what*”, algo que poderá ser um ponto para estudos posteriores.

Também se verificou neste estudo que as perguntas que obtêm respostas mais satisfatórias são as que começam por “*Are*” e “*Where*” possivelmente por exigirem uma resposta simples ou de natureza binária. Perguntas iniciadas com “*Why*” e outras palavras semelhantes que requerem uma resposta explanatória e unívoca são as que obtêm uma resposta menos satisfatória. [Schumaker e Chen, 2010]

### 3.1.2 Emile

Por exemplo, no caso do Agente Emile [Moore e Gibbs, 2002] o seu objetivo era estudar a utilização de Agentes de Conversação para o ensino de teoria social recorrendo à criação de diversas bases de conhecimento de elevada qualidade associadas à personalidade de vários teóricos sociais chave como Marx, Weber, Mead e Foucault. Pretendia-se também que os agentes, para além de serem capazes de responder a pedidos de aprofundar sobre os diversos conceitos, fossem capazes de responder tendo em consideração o nível de conhecimento do utilizador. Outra preocupação da base de conhecimento desenvolvida é indicar o estado atual da discussão sobre os vários conceitos a ensinar expondo também os pontos de vista contraditórios dos vários especialistas.

Foram também implementadas alterações no interface de forma a permitir a gravação das conversas e permitir que o utilizador seja reencaminhado para informação adicional.

O resultado de este estudo permitiu retirar algumas conclusões sobre considerações a ter em futuras implementações de agentes de conversação pedagógicos. De acordo com os questionários realizados neste estudo foi possível deduzir a necessidade da correção ortográfica automática feita por parte do agente, caso contrário este não poderá ser utilizado por estudantes com problemas de escrita.

Uma conclusão positiva de este estudo refere-se ao facto de todos os participantes terem considerado positiva a ajuda do agente na tarefa de aprender sobre teoria social. Os utilizadores referiram que o facto de o Agente ser capaz de explicar conceitos chave e disponibilizar informação adicional que permite aprofundar sobre os vários temas, torna-o uma ferramenta importante para a recolha de informação. Convém também referir o ceticismo dos vários utilizadores em utilizar exclusivamente esta tecnologia sem a utilização de material adicional para aprender sobre um tema. Desta forma, podemos concluir que esta tecnologia é bastante útil quando utilizada como um complemento em tarefas de ensino.

### 3.1.3 Autotutor

Outro exemplo de um Agente com funções pedagógicas é o *Autotutor* implementado para estudar diferentes abordagens na iniciativa ao diálogo. Este agente pretende simular as habilidades sociais que são utilizadas por ambos os intervenientes em uma conversação. A intenção de esta abordagem é permitir a melhoria do desempenho do agente independentemente do domínio de conhecimento em que é utilizado. A iniciativa ao diálogo vai para além do conhecimento linguístico.

Uma das vantagens do *Autotutor* é poder ser utilizado em diversos domínios de ensino que não necessitem de precisão matemática e definições formais. A independência do módulo de conversação faz com que possa ser utilizado em diferentes domínios recorrendo apenas a alterações na base de conhecimento relacionada com o domínio.

Para exemplificar a portabilidade de esta plataforma, foram desenvolvidas em paralelo uma base de conhecimento para o ensino física e outra sobre literatura.

O Auto Tutor dispõe de um mecanismo de análise semântica para atribuir significado às perguntas que lhe são feitas e identificar a qualidade das mesmas. O mecanismo de Análise Semântica Latente (*LSA*) é ideal para representação de enormes quantidades de conhecimento e é capaz de analisar a qualidade de perguntas e as suas respostas com bastante qualidade.

Apesar de este módulo não ser capaz entender o significado de um texto, o seu mecanismo de compressão de conhecimento e identificação de padrões recorrendo a análise estatística é bastante eficiente na identificação de boas e más respostas. Como o *LSA* não analisa a estrutura e a compreensibilidade de um texto e não tem em consideração aspetos sintáticos ou lógicos e torna-se necessário complementa-lo com outros mecanismos tradicionais de

processamento de linguagens naturais para garantir a sua eficiência. [Graesser, et al, 2001]

O *Autotutor* dispõe também de outro módulo que é responsável pelas estratégias utilizadas nas conversações. O estado da conversação é representado em um grafo, isto é uma estrutura de dados que representa as mudanças possíveis nas várias fases de uma conversação. Os vários tipos de interação são classificados como asserções, perguntas, dicas, correções, respostas e as alterações do estado da conversação acontecem de acordo com um conjunto de 15 regras *fuzzy* [Kosko, 1992]. Estas alterações acontecem de acordo com as respostas anteriores e uma avaliação de várias capacidades do utilizador como a sua verbosidade ou iniciativa em conjunto com o tamanho de respostas positivas que foram dadas nesse tópico.

Para testar experimentalmente a eficiência do *Autotutor*, sessenta alunos foram convidados a participar em uma experiência. Cada participante foi avaliado sobre a retenção de informação acerca de três temas de um curso de conhecimento sobre informática. Cada tema foi ensinado de forma diferente para permitir identificar qual o método mais eficiente. Em um dos tópicos de ensino foi possibilitado ao aluno ler e reler sobre o tema antes de ser avaliado. Em outro tópico foi permitido ao aluno ler e mais tarde interagir com o Auto Tutor. No último tópico, seria apenas permitido aos alunos ler uma vez o material de ensino. A cada aluno foi permitido interação com o Auto Tutor de aproximadamente 50 minutos.

Os resultados demonstram mais eficiência na utilização Autotutor do que utilizar os 50 minutos para reler o material de ensino. [Person, et al, 2001]

Apesar de do Autotutor ser capaz de utilizar diferentes abordagens de iniciativa mista tal com descritas por James Allen [Allen 1999] este ainda se encontra longe do cenário ideal de diálogo como acontece naturalmente entre dois intervenientes humanos. De qualquer forma, de acordo com os resultados, podemos considerar o Autotutor um exemplo de como se pode simular a presença de um tutor humano recorrendo a uma abordagem mais activa sobre as várias fases de comunicação. [Lowerse, Graesser e Olney, 2002].

Para além da influência do diálogo do Autotutor na experiência de ensino, outros estudos tentaram verificar a influência da utilização de um Avatar para ajudar a capacidade de controlo da conversação por parte do agente. O agente foi então testado com as quatro seguintes configurações:

- Apenas texto
- Apenas com síntese de voz recorrendo a tecnologia *Microsoft Agent*
- Avatar com síntese de voz e gestos
- Texto e Avatar com síntese de voz e gestos

Os resultados de diversos estudos comprovam o efeito positivo da utilização de agentes no ensino de diversas áreas de conhecimento. Um pormenor interessante é que não foi possí-

vel provar que a utilização de elementos multimédia tenha uma influência significativa na facilitação do ensino. Quanto à influência do diálogo, esta parece tornar-se mais positiva quando se aprofundam os temas de ensino em questão.

Ou seja, podemos concluir que, de facto, a mensagem é o mais importante e, mais uma vez, que a utilização de agentes de conversação pode ter influência positiva no ensino. [Graesser et al. 2003]

## 3.2 O efeito Persona

Para além das questões referidas anteriormente, sabemos também que a presença de um Agente Personificado é, só por si, capaz de melhorar a experiência de aprendizagem remota, é conhecido o efeito *persona*. [Lester et al. 1997].

### 3.2.1 Herman, o inseto

O estudo feito por Lester et al. [Lester et al. 1997] consistiu na utilização de um Agente de Conversação denominado de “Herman o inseto”.

Neste estudo foram convidadas 100 pessoas para testes de utilização de ambas as ferramentas de ensino. Cada pessoa foi convidada a utilizar uma Aplicação de ensino sobre plantas com diferentes implementações de representação do Agente. Foram elaboradas 5 configurações para a utilização do Agente Herman, de forma a poder comparar diferentes níveis de animação e informação dada pelo Agente. Foram utilizados os seguintes tipo de configuração:

- **Completamente Expressivo:** Este agente exhibe todos os tipos de comportamentos comunicativos.
- **Baseado em princípios Animado/Verbal:** Este tipo de agente só fornece informação baseada em princípios, acompanhada de animação visual e descrição sonora.
- **Baseado em princípios Verbal:** Este agente é semelhante ao descrito anteriormente mas não fornece informação visual.
- **Tarefas Específicas Verbal:** Este agente só fornece informação verbal relacionada com tarefas em específico.
- **Mudo:** Este agente não fornece conselhos sobre componentes da planta.

Os utilizadores responderam a um questionário sobre o tema em causa, antes de procederem à utilização das ferramentas de ensino. Desta forma foi possível dividir os utilizadores em diferentes grupos de acordo com o seu nível de conhecimento sobre plantas. Após várias sessões de interação com as diferentes configurações da plataforma, os utilizadores foram convidados a responder a outro questionário de forma a avaliar a retenção de informação e também aspetos como a satisfação obtida a partir da interação com o sistema.

Foi possível verificar o impacto do efeito *Persona* nas diferentes configurações utilizadas experimentalmente. De acordo com os estudantes, mesmo a versão menos interativa, foi capaz de provocar melhorias na experiência de aprendizagem. Este estudo leva-nos a considerar a possibilidade a presença de uma simples representação visual, sem funções práticas, permitir, só por si, a melhoria da experiência geral dos utilizadores. De acordo com o autor deste estudo, não é previsível que as respostas dos utilizadores quanto à avaliação da ferramenta tenham sido influenciadas, devido ao uso de certas precauções como o anonimato e também por as respostas se encontrarem de acordo com os resultados obtidos na avaliação feita ao conhecimento retido.

Para além das implicações positivas de este estudo, é importante considerar que os agentes de conversação também podem ter influências negativas, caso se demonstrem excessivamente intrusivos e proactivos.

É possível então concluir que no caso da utilização de agentes de conversação, se deve considerar sempre a utilização de uma representação visual animada mas que no caso de se utilizar uma estratégia de ensino proactiva, será necessário uma implementação mais cuidada. Apesar de este estudo não permitir concluir acerca da influência da utilização de esta tecnologia a longo prazo, existe a possibilidade de o efeito de motivação proporcionado pela interação com o agente se acentuar com um uso mais prolongado.

Este estudo foi efetuado, recorrendo a um grupo de 100 estudantes, utilizando apenas uma das bases de conhecimento desenvolvidas. Espera-se que outros estudos comprovem estes resultados recorrendo a diversas bases de conhecimento e grupos etários distintos.

Outros estudos referem também a importância de todos os elementos sensoriais como uma possibilidade de aumentar a credibilidade do agente atribuindo principal importância aos elementos visuais e sonoros [Thalmann N. Magnenat *et al.* 2005] [Koda e Maes 1996].

Pretende-se então analisar individualmente a influência dos elementos visuais e sonoros no efeito *persona* de forma a compreender qual o peso de cada um destes elementos na melhoria da experiência de aprendizagem.

### 3.2.2 Freudbot e Piagetbot

Existem diversos estudos sobre o efeito *Persona* quando utilizamos Agentes em tarefas pedagógica, contudo em parte os seus resultados são contraditórios. Centrados na ideia do efeito *Persona*, [Heller et Al.] decidiram comprovar este efeito utilizando um Agente que representado por uma figura familiar. Neste caso, optaram por utilizar a imagem Sigmund Freud que é figura histórica conhecida. Este agente, o Freudbot, seria responsável por disponibilizar informação autobiográfica durante uma sessão de conversação.

Apesar de os resultados comprovarem que os utilizadores consideraram a experiência positiva e bastante intuitiva. O formato de aprendizagem através de conversação parece de facto adequado para aprender informação biográfica sobre uma figura histórica. Contudo,



por falta de dados que permitissem a comparação, este estudo foi incapaz de demonstrar e quantificar a vantagem da aprendizagem feita por meio de conversação em vez de se utilizar consulta direta.

Para obter resultados mais objetivos foi realizado por Bob Heller e Mike Procter uma experiência semelhante à do Freudbot. Desta vez, foi concebido um agente semelhante ao Freudbot, mas com a informação biográfica relativa a Jean Piaget, o Piagetbot.

Para além do desenvolvimento do agente, foi disponibilizada, em paralelo, a mesma informação através de uma página web contendo apenas texto. A informação disponibilizada pelo Agente seria dada na primeira pessoa ao contrário da informação disponível através da página que se referia a Piaget na terceira pessoa.

Estas duas possibilidades serviram para tentar comprovar a vantagem da aprendizagem através de conversação em relação a um método, que apesar de interativo, consiste na consulta direta.

Surpreendentemente, os resultados de este estudo contrastam com os referidos por Lester et al [Lester *et al.* 1997], pois as respostas dadas pelos participantes indicam a sua preferência pelo método não conversacional por ser mais prático. Uma das possibilidades de se ter verificado este resultado pode estar relacionada com o formato da experiência.

Como os participantes haviam sido informados que iriam responder a um teste onde seriam avaliados sobre o seu conhecimento sobre Jean Piaget, estavam obviamente interessados em recolher o máximo de factos importantes sobre o assunto em um curto espaço de tempo. Neste contexto, é possível compreender a vantagem de um método que permite aceder à informação de forma direta mas não é possível concluir sobre a possibilidade de ambos os métodos serem utilizados em um contexto sem restrições temporais.

### **3.2.3 Outros estudos sobre a influência da representação visual**

É possível encontrar uma análise feita ao diferente tipo de representações visuais de agentes. Os resultados demonstram que não existe diferença quanto ao sexo da representação do agente, porém parece existir um benefício positivo quanto à utilização de uma personagem realista em vez de caricaturada.

Estes resultados foram comprovados por Tomoko Koda e Pattie Maes através de uma de diversas experiências realizadas para avaliar a influência de diversos aspetos relacionados exclusivamente com a aparência do agente de conversação, nomeadamente, o sexo, o aspeto humano e realismo. Os resultados experimentais demonstraram que não existe grande influência quanto à escolha de sexo do agente.

Quanto ao aspeto humano, foi possível comprovar que um agente de conversação com aparência humana é percecionado como mais inteligente, contudo, menos cativante que um agente representado por um animal, neste caso, um cão.

Na experiência efetuada para comparar o impacto de diferentes níveis de realismo da cara do agente, os resultados indicam melhor desempenho de qualquer cara realista, independentemente do nível de detalhe, quando comparadas com um agente representado por uma caricatura ou uma personagem como um “*smiley*”, ou seja, uma aparência não realista

Convém salientar que este estudo é resultado de um teste feito com agentes que desempenham a tarefa de jogar póquer. Desta forma não podemos concluir com certeza que se obtém um aumento de credibilidade utilizando estas técnicas em um Agente de Conversação. [Koda e Maes 1996]

Convém notar que os vários estudos sobre este tema chegam a ser contraditórios e torna-se difícil concluir objetivamente sobre a influência do efeito persona. Grande parte destas contradições assentam sobre o facto de as experiências serem realizadas em circunstâncias bastante diferentes o que impossibilita estabelecer uma relação entre os diversos estudos. Por exemplo,

Beun, Vos e Witteman referem a melhoria da capacidade de memória do utilizador ao utilizarem uma imagem para representar o agente em uma simples tarefa de conversação. Contudo, este estudo conclui também que a aparência humana do agente, apesar de ter influência na percepção do utilizador, não produz melhorias quanto à capacidade de memorização. [Beun, Vos, Witteman 2003]

Estudos anteriores revelam resultados semelhantes em contextos diferentes. O estudo efetuado por King e Ohya [King e Ohya 1996] centra-se exatamente sobre a influência das características antropomórficas quando utilizamos a representação visual de um agente de conversação. Este estudo conseguiu relacionar o aspeto humano de um agente com a percepção da sua inteligência.

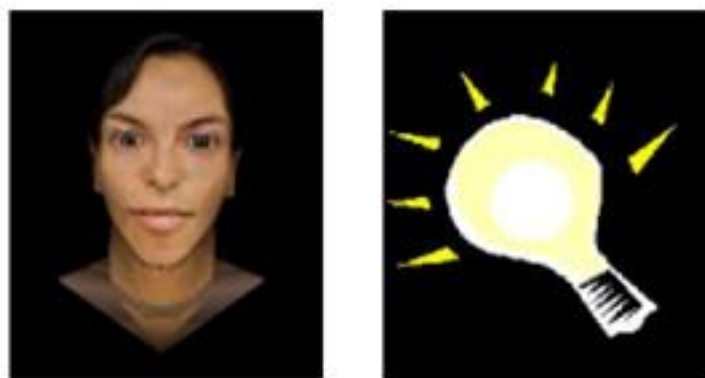
Recorrendo à utilização de representações de diferentes objetos como uma lâmpada e utilizando vários níveis de realismo foi possível concluir acerca das vantagens da utilização de um agente animado com características humanas o mais realista possível. Para além de um agente com aspeto humano ser considerado mais inteligente do que todas as outras representações, também foi possível verificar que no caso de esta personagem ser animada, exibindo um simples piscar de olhos, obtém resultados ainda mais favoráveis.

Trinta e nove alunos participaram na fase experimental do estudo e foram-lhes atribuídas tarefas de forma aleatória. Uma das tarefas consistia em efetuar o planeamento de uma viagem e neste caso a presença do Agente de Conversação permitiria uma troca de opiniões sobre a importância dos artigos a escolher. Essencialmente, os participantes tinham de escolher 6 artigos de 12 para levar para uma viagem específica e ordena-los de acordo com a sua importância. Posteriormente, em forma de conversação, o agente indicaria então qual a sua opinião sobre a escolha do utilizador e seria permitido ao utilizador efetuar alterações à sua escolha.

A outra tarefa consistia na edição de um texto e neste caso o agente poderia ser usado para obter referência das funções para utilizar no programa de edição de texto. Os utilizadores teriam de efetuar várias alterações a um texto. Apesar de ser mostrado aos utilizadores um vídeo a descrever as várias funções disponíveis no editor de texto, a quantidade de funções disponíveis no editor e o número de comandos a executar foram planeados de forma a fazer com que a ajuda do Agente de Conversação se tornasse necessária para completar a tarefa.

No fim de cada tarefa, os participantes foram convidados a responder a um questionário acerca das dificuldades no desempenho da mesma e da influência dos Agentes de conversação.

Entre as várias sessões de teste foram alteradas as representações do Agente de conversação utilizando uma versão animada, outra sem movimento e uma representação icónica. A ordem em que foram executadas as tarefas também foi alterada.



*Ilustração 4. Aparência do Agente humana e do agente icónico [Xiao, Stasko e Catrambone 2002]*

Para além de um monitor a auxiliar o participante, foi utilizado um monitor que garantiu a apresentação de respostas válidas, escolhendo-as a partir de um conjunto.

A escolha das duas tarefas foi feita de modo a poder testar-se a influência vários tipos de agente quando utilizados para auxiliar em tarefas de diferente objetividade. Neste caso, a tarefa das viagens corresponde permite alguma subjetividade na escolha das respostas enquanto a tarefa de edição de texto é bastante objetiva pois é apresentado ao utilizador qual deve ser o resultado final. Dada a diferença das tarefas, tal como já havia sido referido, a maneira como o Agente se dispões para ajudar também muda sendo que na tarefa de edição este encontra-se sempre disponível e apenas reage quando pedido pelo utilizador, enquanto na tarefa de viagem este interfere sempre independentemente da ação do utilizador.

Os resultados obtidos experimentalmente permitem concluir que, apesar de não se ter verificado melhorias no tempo de execução das tarefas de teste, foi possível comprovar que na tarefa da viagem, os participantes tiveram em consideração as alterações propostas pelo agente durante as tarefas independentemente do tipo de agente utilizado.

A partir das respostas obtidas por questionário sabemos também que, no geral, os participantes consideraram a ajuda do agente como sendo positiva. Por outro lado, também foi possível verificar que a satisfação depende e da tarefa a desempenhar e neste caso o agente foi considerado mais intrusivo ao ajudar na tarefa de escolher as viagens. Nesta tarefa observou-se que foram poucos os participantes que responderam a interações do agente sem ser através de respostas curtas.

Um dos dados que se revela interessante é o facto de o agente das tarefas de viagens ser considerado com mais personalidade o que resulta de este emitir opiniões em vez de indicações concretas como no caso da tarefa de edição.

Outro facto que foi possível observar a partir da resposta dos utilizadores é a forma como estes se referem aos agentes. É curioso observar que um quarto dos utilizadores referiu-se ao agente utilizando pronomes pessoais, ou seja, como uma pessoa, o que salienta a noção de como por vezes os agentes são entendidos como atores sociais [Nass, Steuer e Tauber 1994]. Mais curioso ainda é o facto de a grande maioria dos utilizadores que se referiram aos agentes como pessoas, serem do sexo feminino. [Xiao, Stasko e Catrambone 2002].

Quanto à influência específica da aparência do Agente, é possível verificar que os vários estudos produziram alguns resultados contraditórios e poucos resultados completamente inequívocos. Apesar de a exibição de uma face ser considerada divertida, esta pode diminuir a concentração dos utilizadores ao atrair a atenção com facilidade. Outras imagens que captem menos a atenção do utilizador, como uma seta, podem então ser mais uteis em algumas situações [Takeuchi e Naito 1994]

Apesar de ser considerado improvável que implementações que tenham em conta apenas a melhoria estética proporcionada pela representação visual do agente sejam capazes de proporcionar benefícios [Cassell 2000], não nos podemos esquecer que existem diversas variáveis como a tarefa a desempenhar ou o domínio de conhecimento a utilizar e que no caso de verificação experimental, torna-se difícil obter resultados comparáveis.

### **3.3 Agentes de Conversação com representação emocional**

O uso de representações emocionais, que podem ser efetuadas através de diversas alterações, seja no discurso, na representação visual do agente ou até no seu tom de voz, tem recebido alguma atenção em pesquisa feita tanto no campo de Ambientes Virtuais como de Agentes de Conversação.

Existe a tentativa de aproximar os dois campos de estudo de forma tornar a integração de Agentes de Conversação em Ambientes Virtuais mais natural e detalhada, resultado em uma experiência mais imersiva e credível para o utilizador. [Thalmann N. Magnenat et al 2005]

Esta abordagem segue também a ideia descrita por Thalmann sobre a influência do efeito Persona. Desta forma, a capacidade do Agente ser capaz de demonstrar reações emocionais adequadas às situações, tal com um humano, pode ser analisada como forma de melhorar a sua eficiência.

### **3.3.1 Representação de emoções**

Existem diversos estudos centrados na representação e variação das emoções do agente recorrendo à extensão do AIML [Galvão, et al, 2004][Fitrianie, et al, 2003] [Thalmann e Kshirsagar 2002]. Foi também possível verificar experimentalmente a vantagem da utilização da representação emocional em um ambiente de conversação virtual [Vrajitoru 2006] [Cho 2007].

As emoções são estados mentais que ocorrem momentaneamente resultantes de impulsos e não através de pensamentos conscientes. Existem diferentes modelos para representar as diversas emoções sendo os mais utilizados o modelo definido por Ekman [Ekman e Rosenberg 2005] e também o modelo criado por Ortony, Clore e Collins [Ortony, Clore e Collins 1988].

No caso do modelo Ekman, são as seis emoções que podem ser expressas através da face em qualquer cultura, nomeadamente, felicidade, tristeza, raiva, medo, repugnância e surpresa. Este modelo tem a vantagem de poder ser facilmente representado recorrendo em um personagem virtual e, tal como proposto por Ekman, é possível obter outras expressões também facilmente identificáveis, a partir da sobreposição de diferentes características das seis emoções base [Ekman e Rosenberg 2005].



*Ilustração 5. Emoções básicas (repugnância, felicidade, tristeza, raiva, medo, e surpresa)  
[Vinciarelli, Pantic e Bourlard 2009]*

O modelo OCC considera que as emoções são adequadas a eventos, agentes e objetos. Essas reações poder ser separadas em três categorias distintas de emoções: Satisfeito/Insatisfeito, Aprovado/Desaprovado, Gostar/Não Gostar.

De forma a poder utilizar o Modelo OCC em conjunto as emoções definidas por Ekman, é possível reduzir as 22 emoções descritas no Modelo OCC às definidas por Ekman. Assim sendo torna-se possível utilizar a tabela de alterações definida no Modelo OCC em conjunto com um número reduzido de expressões que podem ser facilmente representadas. As emoções Medo e Repulsa foram adicionadas ao Modelo OCC este não ter emoções equivalentes e haver interesse em representá-las. [Thalmann e Kshirsagar 2002]

<b>Ekman</b>	<b>Modelo OCC</b>
Felicidade	Feliz-por, Regozijo, Felicidade, Orgulho, Admiração, Gratificação, Gratidão
Tristeza	Ressentimento, Pena, Vergonha, Dificuldade, Remorsos
Raiva	Raiva, Reprovação, Ódio
Surpresa	Surpresa
Medo	Medo, Medo-confirmado
Repulsa	Repulsa

*Tabela 6. Correspondência de emoções básicas definidas por Ekman e pelo Modelo OCC  
[Thalmann e Kshirsagar 2002]*

Um dos problemas do modelo OCC para a implementação de um sistema de gestão de emoções, deriva do facto de se poder encontrar ambiguidades. Este problema foi identificado e foram propostas alterações ao modelo de forma a permitir uma implementação consistente. [Steunebrink, Dastani e Meyer 2009]

De forma às alterações emocionais poderem ocorrer de acordo com o Modelo OCC, é também necessário considerar o estado atual do agente. O estado do agente é definido pelo seu humor e pela sua personalidade. A representação da personalidade é normalmente feita de acordo com o modelo *Factor Five* que a divide em cinco fatores. Desta forma podemos definir o tipo de personalidade do nosso agente, atribuindo um valor percentual a cada um dos fatores.

Como o humor é também um reflexo da personalidade do Agente, que também teria de ser definida, a informação disponível não é suficiente para permitir uma implementação adequada. Desta forma, opta-se por uma personalidade neutra e o humor é apenas representado por positivo e negativo.

A informação associada ao estado emocional atual do agente é processada recorrendo a um grafo *Bayesiano* que representa os vários estados e sucessivas transições, tal como descritas pelo modelo OCC. Foi utilizada esta estrutura de dados por permitir a utilização de incerteza, que é uma característica indissociável do comportamento humano. Essencialmente um grafo Bayesiano é um grafo direcionado sem ciclos, onde cada nó representa o valor de um de dois estados mutuamente exclusivos. Cada aresta do grafo representa a influência do nó pai sobre os seus sucessores. O estado dos nós filhos é determinado de acordo com uma tabela de condições e probabilidades e o valor do seu nó pai. A tabela que define os valores das transições é implementada na definição do grafo e os seus valores foram escolhidos de forma intuitiva, e tal como sugere o autor do estudo pode ser uma área que necessite de uma pesquisa mais extensa.

Outros trabalhos [Ball e Breese 1998] já haviam utilizado grafos *Bayesianos* para representar um modelo com personalidade e emoção. O trabalho desenvolvido por Thalmann e Kshirsagar [Thalmann e Kshirsagar 2002] diferencia-se por acrescentar a influência do humor ao modelo.

As emoções associadas a cada resposta podem ser guardadas na base do conhecimento do Agente recorrendo à adição de novas *tags* no AIML. Esta extensão do AIML permite atribuir uma intensidade às seis emoções.

A principal desvantagem de este método é o facto de implicar uma edição extensa da base de conhecimento e, mesmo assim, a informação emocional é ditada pela pessoa responsável por escolher os valores associados a cada emoção.

Outra abordagem possível, a utilização de informação externa que permita associar palavras-chave a emoções. Através do uso de uma base de dados de senso-comum como, por exemplo, a *Open Mind Common Sense Knowledge Base* é possível quantificar as emoções

existentes em um texto. Este mecanismo identifica a palavra principal de um texto e, através da base de dados, obtém um sêxtuplo que representa as várias emoções (felicidade, tristeza, raiva, medo e surpresa) associadas a essa palavra. Para se obter um valor mais preciso, são analisadas recursivamente frases que se encontram normalmente associadas à palavra inicial são adicionadas as emoções ainda não detetadas. O resultado é atenuado pelo fator de propagação, referido no exemplo como a variável  $d$ . [Liu, Lieberman e Selker 2003]

*“Algo excitante é simultaneamente alegre e surpreendente.*

*(passagem: 1: excitante [1,0,0,0,0,1])*

*Montanhas-russas são excitantes.*

*(assumir  $d=0.5$  passagem: 2: montanha-russa [0.5,0,0,0,0,0.5])*

*As montanhas russas normalmente encontram-se em parques de diversões.*

*(passagem 3: parque de diversões: [0.25,0,0,0,0,0.25])”*

[Liu, Lieberman e Selker 2003 tradução nossa]

Esta técnica foi testada utilizando uma aplicação denominada *EmpathyBuddy*. Os utilizadores foram convidados a enviar uma mensagem de correio eletrónico a um amigo através da aplicação que é responsável por identificar e exibir a emoção predominante enquanto o texto é escrito.

Os dados obtidos a partir da utilização experimental, são curiosos porque este método revelou-se menos emocionante para a maioria dos utilizadores quando comparado com uma implementação que exibia emoções de forma aleatória. Contudo, o dado mais importante e motivador deste estudo é o facto da utilização do método descrito ser considerada pelos utilizadores como a implementação com a interpretação mais inteligente.

Outra técnica para identificar as emoções associadas a uma frase é descrita por Xu Zhe e Anthony Boucouvalas [Zhe e Boucouvalas 2006] que conseguiram desenvolver uma aplicação que permite a extração de essa informação em tempo real. De forma a ser possível a identificação de palavras associadas a emoções, teve de ser desenvolvido um dicionário com a caracterização de 16400 palavras. Uma das preocupações que surge nas conclusões deste estudo relaciona-se com o problema dos erros ortográficos e gramáticas, algo que já havia sido referido [Schumaker, Ginsburg, Chen e Liu 2006]

Recorrendo à utilização de ambas técnicas em conjunto seria possível obter informação mais precisa e completa sobre o impacto emocional que resulta da frase em questão.

Existem também tipos de interações e ações não-verbais que podem ajudar a melhorar a credibilidade do Agente de Conversação como por exemplo o movimento da cabeça e dos olhos que são elementos importantes durante a comunicação. O movimento das sobrancelhas deve também ser considerado por ser uma parte importante na representação de emoções. [Cho 2007]



De acordo com um questionário realizado a um grupo de utilizadores, foi possível aumentar o sucesso do agente de conversação em relação à implementação tradicional do Alice-Bot [Cho 2007]. Convém salientar que, neste exemplo, a instabilidade emocional do agente de conversação foi considerada como um fator negativo por um dos utilizadores.

A sincronização dos movimentos faciais do agente, para representar a fala e as expressões emocionais, é necessária para aumentar o realismo do agente de conversação. Uma abordagem possível é efetuar a animação facial durante o tempo da tarefa de fala. No caso de [Thalmann e Kshirsagar 2002] optaram pela utilização de técnicas estudadas anteriormente [Kshirsaga, Molet e Thalmann 2001]. Para além da animação da fala e das emoções também é possível alterar diversos parâmetros como movimentos da cabeça e piscar de olhos.

Existem questões relacionadas com a credibilidade do agente que ainda não foram analisadas. Isto é, para um máquina conseguir simular o comportamento humano, é também necessário replicar algumas das suas limitações humanas.

As respostas dadas por um agente de conversação são determinadas pelo conteúdo da sua base de conhecimentos e são processadas bastante rápido quando comparadas com o tempo médio de uma pessoa a escrever em um teclado. Assim sendo, torna-se necessário introduzir uma latência nas ações do agente de conversação e a possibilidade de forçar erros de escrita. Estas questões foram obviamente consideradas e estudadas por alguns dos trabalhos candidatos a ganhar o Prémio Loebner.

Outra diferença entre os humanos e as máquinas, para além da questão da velocidade processamento, relaciona-se com a precisão. Convém ter em conta que no caso de perguntarmos a um agente uma questão aritmética, este revelará o resultado com uma precisão que não consegue ser calculada facilmente por um humano. Desta forma, torna-se necessário, mais uma vez, implementar mecanismos que limitem a capacidade dos Agentes de forma a agirem de uma forma humana.

Em relação à quantidade de informação disponível pelo agente, convém lembrar que, de acordo com os resultados da primeira edição do Prémio Loebner, os humanos são capazes de exibir conhecimentos tão profundos sobre um tema que não existe necessidade de restringir a capacidade do Agente.

Estas técnicas são apenas importantes se pretendermos desenvolver um agente com o objetivo de encobrir a sua identidade. No caso de o agente ser concebido para desempenhar um papel de tutor, estas técnicas não devem ser consideradas.

No caso de estarmos a avaliar um Agente de Conversação com funções pedagógicas, o principal objetivo do Agente é facilitar o acesso a informação correta e tornar a tarefa mais agradável e intuitiva.

De forma a permitir a consulta sobre informação que não está disponível na base de conhecimento do agente, o ideal seria a utilização de consultas a fontes exteriores, como a Internet, ou recursos disponíveis em rede.

Uma abordagem a este problema consiste na utilização de informação disponível na Internet em conjunto com ferramentas como o *Wordnet*. A partir da identificação das palavras mais importantes de uma pergunta e através do relacionamento semântico da informação obtida é possível obter uma resposta. De forma a melhorar este processo, a base de conhecimento *Wordnet* é utilizada para obter sinónimos das palavras identificadas. [Yang et al, 2006]

Convém salientar que este método é responsável pela geração de uma resposta a partir da informação como esta se encontra disponível, com mais frequência na Internet e nem sempre é possível a obtenção de uma resposta. A informação pode não estar disponível ou não ser possível obter informação a partir da formulação da pergunta.

Como este método é complementar aos utilizados pelos Agentes de Conversação e pode ser usado como uma forma de expandir a capacidade do Agente sendo apenas utilizado em circunstâncias específicas.

### 3.4 Conclusão

Em este capítulo foi possível concluir acerca de diversas características capazes de influenciar o desempenho de várias ferramentas que utilizam agentes de conversação.

Para além da influência da linguagem utilizada pelo agente durante a conversação [Schumaker e Chen, 2010], também existem estudos que concluem acerca da utilização de representações visuais para o agente de conversação, o denominado efeito Persona [Lester *et al.* 1997]. Este efeito revela-se mais influente quando se utiliza uma aparência mais humana e realista para caracterizar o agente de conversação.

Em relação à utilização de emoções para atribuir ao agente características ainda mais humanas, a sua utilização é considerada positiva por vários estudos. Em estes estudos o estado emocional do agente serve para alterar a representação visual do agente mas também para apresentar respostas mais adequadas. [Cho 2007]

## Capítulo 4

# 4. Plataforma desenvolvida

Antes de iniciar o desenvolvimento da plataforma foi necessário identificar um conjunto de funcionalidades para se poder proceder à escolha da tecnologia adequada. Esta plataforma pretende diferenciar-se dos trabalhos anteriormente descritos por permitir a interação e avaliação, de diferentes aspetos de um agente, de uma forma remota.

### 4.1 Especificação

Foi então necessário escolher tecnologia que permitisse a interação com Agentes de Conversação através de protocolos de *Instant Messaging*, considerando os elementos multimédia que serão adicionados como a visualização de *Avatares* e a utilização de emoções. Outro aspeto importante da plataforma que tem de ser considerado é a forma como são guardados os dados, respetivos à sua utilização, para se poder posteriormente proceder à sua avaliação.

#### 4.1.1 Servidor de Conversação

De forma a possibilitar o acesso remoto, pretende-se implementar esta interface, de forma a poder ser executado em um *browser*, recorrendo a tecnologia como o Flash ou PHP. Os serviços utilizados para a comunicação e para a tecnologia de agentes devem ser convergentes, usando, por exemplo, uma especificação como o *Extended Markup Language (XML)*.

Neste caso, a configuração ideal parece ser o uso de *XMPP* [*Jabber Software Foundation*] para o servidor de conversação, que é desenvolvido sobre protocolo que utiliza o XML e possui bastantes extensões como a integração com tecnologia de ligação, conhecida como *gateways*, para interagir com outros protocolos de *Instant Messaging*.

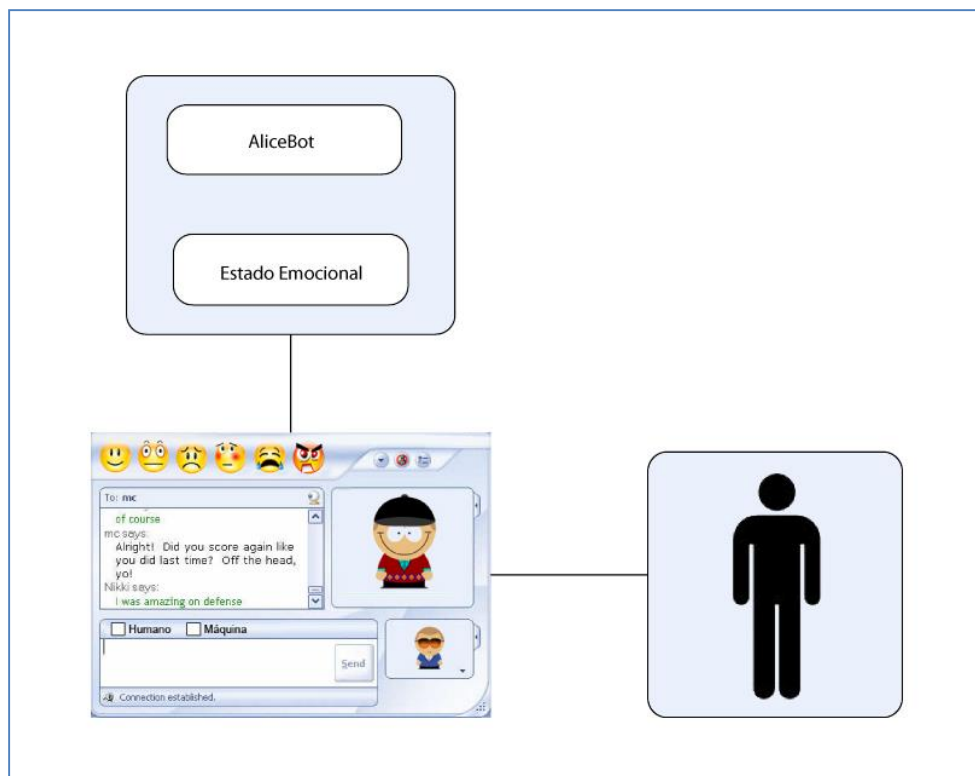
Uma das vantagens da utilização de tecnologia como o *Adobe Flash* é permitir bastante flexibilidade, tanto na implementação do interface de conversação como a visualização dos estados emocionais.

### 4.1.2 Agente de Conversação

Para a tecnologia de agentes de conversação, tal como referido anteriormente, pretende-se utilizar uma variante do conhecido *Elizabot*, o *AliceBot*, por utilizar uma base de conhecimento em AIML que é também desenvolvido em XML.

Como opções para a implementação do agente de conversação, podemos utilizar o *Persona-Bot* que apresenta como principal dificuldade a obtenção de esta tecnologia ou, como alternativa, implementar uma versão simplificada da tecnologia referida anteriormente.

Em ambos os casos, o conteúdo da base de conhecimento terá de ser alterado em quase toda a sua extensão, de forma a se adicionar a informação relativa aos estados emocionais. Poderemos então usar a estrutura referida anteriormente ou então substituí-la por algo mais simples, como o uso de sequência caracteres utilizados normalmente na comunicação escrita para representar emoções, os "smilies" ou "emoticons".



*Ilustração 6. Exemplo de conversação entre agente e utilizador*

No caso de se recorrer à segunda hipótese, podemos utilizar um servidor como o *XMPP Chatbot*, que se encontra disponível como projeto de código aberto. Neste caso, teremos de optar entre alterar o funcionamento do servidor para calcular as alterações das emoções de acordo com o desfecho das conversas, ou atribuir estas funcionalidades ao cliente, de forma simplificada. Teria também de ser utilizado um servidor como o *MySQL* para guardar as variações de emoções, desfecho das interações, juntamente com todos os parâmetros que possam interessar para posterior análise.

O Xmp Chat Bot é um script em *Perl* que serve de ligação entre um servidor XMPP e um agente de conversação Pandorabot. A instalação e configuração são bastante simples. Consequentemente, os parâmetros de configuração do script são apenas os dados para efectuar a autenticação das contas *XMPP* e *Pandorabots*.

Os agentes de conversação *Pandorabot* são implementações do *Alicebot* que são configuráveis a partir de um interface web. Assim sendo, podemos definir características do agente de conversação como a idade, nome ou nacionalidade e também adicionar diferentes bases de conhecimento em *AIML* recorrendo ao uso de um browser.

Um dos problemas que surgiu devido à utilização de *AIML*, é a falta de bases de conhecimento que utilizem a língua Portuguesa. De forma a permitir que os testes sejam feitos recorrendo à língua Portuguesa, torna-se então necessário escolher uma base de conhecimento para proceder à sua tradução.

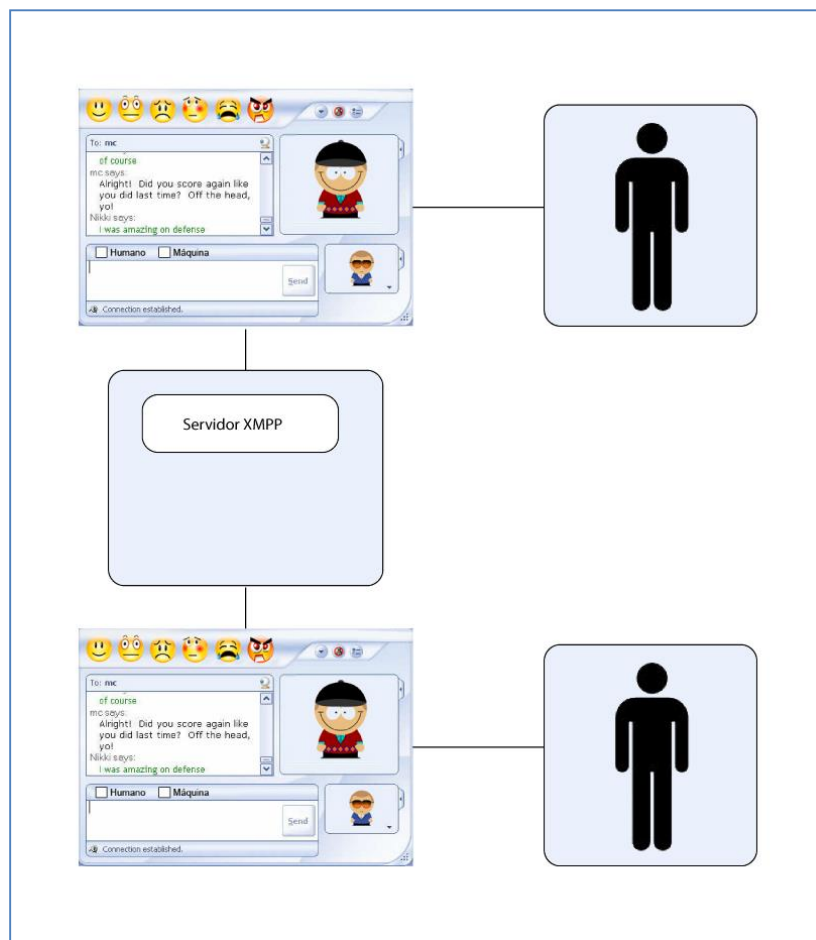


Ilustração 7. Exemplo de conversação entre dois utilizadores

### 4.1.3 Funcionalidades a Implementar

De acordo com o objetivo do trabalho, é necessário escolher um conjunto de funcionalidades que apenas será utilizado em uma segunda fase de testes. Estas funcionalidades têm de

ser identificadas para poderem ser implementadas de forma a poderem ser configuradas para diferentes tipos de sessão de teste.

Para além da representação visual dos intervenientes e respetivas alterações emocionais, que são duas funcionalidades distintas, torna-se necessário implementar outras funcionalidades.

Como o tempo de resposta do agente de conversação depende apenas da comunicação com o serviço *Pandorabots*, a resposta será exibida ao utilizador de forma quase instantânea, independentemente do tamanho da resposta. Desta forma, para simular a demora de uma mensagem ser escrita por um humano em um computador, pode-se introduzir uma latência que esteja dependente do número de caracteres da resposta.

Para replicar as sessões tal como avaliadas no Prémio Loebner torna-se necessário controlar o fluxo de diálogo para que o diálogo seja sequencial e composto por apenas um parágrafo. [*Loebner Prize*]

### 4.1.4 Recolha e armazenamento de Informação

Para além da informação que será recolhida durante as sessões de conversação, torna-se necessário definir um questionário que será efetuado aos utilizadores no fim da mesma. O objetivo do questionário é permitir quantificar o conhecimento do utilizador sobre a tecnologia a ser testada. Outro objetivo do questionário é saber qual a influência das funcionalidades durante as conversações e descobrir quais foram os fatores fundamentais de sucesso ou insucesso do agente de conversação.

A informação relativa à interação do utilizador com a interface e também às várias sessões de conversação devem também ser guardadas para posterior análise. Como a informação, relativa aos utilizadores, se encontra guardada em uma base de dados relacional *MySQL*, podemos expandi-la para guardar os resultados das interações resultantes da fase de teste. Também será possível utilizar esta base de dados de forma a guardar um conjunto de configurações que caracterizem as diferentes fases de teste através de pormenores como a duração de cada sessão e o conjunto de funcionalidades a utilizar.

## 4.2 Implementação

### 4.2.1 Servidor de Conversação

Para o desenvolvimento da Plataforma de Avaliação, foi utilizado o servidor XMPP *Openfire* por preencher os requisitos descritos anteriormente, ou seja, a possibilidade de integração com o Flash através do uso de *sockets XML*.

Durante a instalação do servidor, foi necessário proceder à criação da respetiva base de dados MySQL, para esta ser usada em alternativa à base de dados interna do servidor *Openfire*.

Esta tarefa é simplificada pelo facto de podermos criar um grupo de utilizadores no servidor de XMPP que corresponde aos agentes de conversação. Também é possível configurar o servidor para que todos os utilizadores novos, após o registo, irão adicionar automaticamente à sua lista de contactos, todos os elementos do grupo de agentes.

Para além dos agentes de conversação, torna-se necessário gerir a presença dos vários utilizadores. Isto é, temos de considerar a disponibilidade dos vários utilizadores quando se encontrem na plataforma de teste. A opção escolhida para esta gestão foi a utilização de uma sala de conversação. Isto porque não existe necessidade de o utilizador ter acesso a informação relativa à presença de outros utilizadores que não estão disponíveis.

Quando um utilizador entra na sala aplicação tenta procurar um utilizador que esteja disponível. No caso de todas as respostas serem negativas, o utilizador será notificado e terá de aguardar pela presença de novos utilizadores.

Para além da gestão dos privilégios dos utilizadores e salas de conferência, o servidor *XMPP* permite guardar registo de todas as conversações.

### 4.2.2 Agente de Conversação

Após a criação de uma conta no serviço *Pandorabots*, torna-se disponível o acesso à configuração de um agente de conversação. Através de este serviço, é possível também alterar a base de dados de conhecimento do agente através do carregamento de ficheiros *AIML*. Podemos alterar outros dados do nosso agente, como o nome, a idade, etc. Depois de o agente estar configurado e devidamente publicado, passa a ser possível testá-lo a partir do interface do serviço ou através de pedidos *HTTP* feitos para o servidor *Pandorabots*.



Ilustração 8. Interface de administração de agentes *Pandorabots* [*Pandorabots*]

Para o desenvolvimento de este trabalho, tal como descrito anteriormente, a opção escolhida foi a utilização de um *script* em *Perl* [*Perl*] que serve como ligação entre uma conta *XMPP* e um agente *Pandorabots*. Será possível conhecer a disponibilidade de todos os

agentes de conversação pois poderão ser adicionados automaticamente à lista de contactos de todos utilizadores.

Neste caso, torna-se necessário a utilização de vários agentes em simultâneo para ser possível comparar a eficiência de um Agente dotado de uma base de conhecimento geral e outro especializado em discutir um tópico específico como, por exemplo, Desporto.

### 4.2.3 Cliente de Conversação

Para além das funcionalidades já existentes no servidor de conversação, torna-se necessário implementar no cliente de conversação as ações adequadas às várias fases de cada sessão.

Considerando que o cliente de conversação foi completamente desenvolvido no âmbito de este trabalho, foi necessário implementar todas as funcionalidades básicas de uma ferramenta de conversação. Esta tarefa foi facilitada através do uso das classes *XIFF*.

Tal como referido anteriormente, a biblioteca *XIFF* serve como elo de ligação entre o *Flash* e o *XMPP* enviado pelo servidor. Esta é responsável por alertar o *Flash* de todas as ocorrências em uma ligação ao servidor, desde a autenticação do utilizador à gestão de mensagens, e traduzir essa informação para objetos que podem ser acedidos pelo mesmo.



*Ilustração 9. Interface de autenticação ou registo do utilizador*

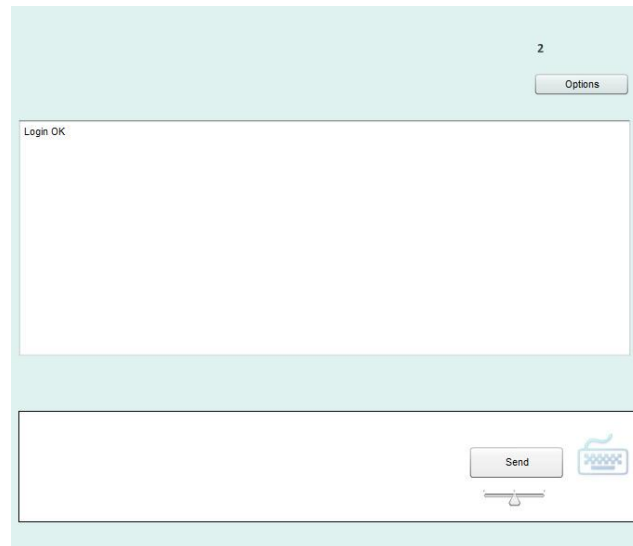
Para ser possível aceder à ferramenta, os utilizadores têm de efetuar o registo de forma a ser possível identificar algumas características simples como a idade e sexo.

#### 4.2.3.1 Tipos de sessão

Uma das características principais da plataforma desenvolvida é o facto de permitir a configuração de várias características que determinam a aparência e funcionamento do interface de utilização. Isto é, o administrador do sistema pode escolher qual o número de sessões experimentais que irão decorrer, qual a sua duração e determinar se o parceiro de cada conversação será um humano ou um agente. Outra característica que pode ser alterada é a aparência do interface dentro de um conjunto possível.

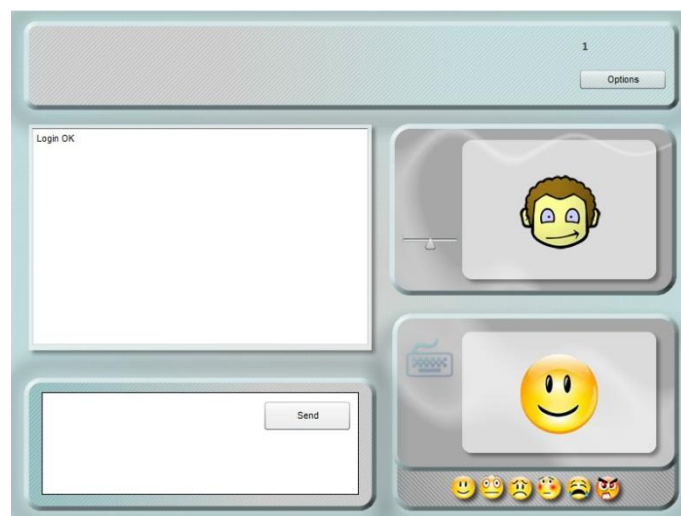


A primeira opção é representação mais parecida com o teste inicialmente proposto por Turing. A ausência de representações do Agente de Conversação e de outros indicadores do estado emocional da conversa faz com que o centro da atenção seja a conversa. Desta forma, o utilizador pode concentra-se em analisar apenas a capacidade do diálogo do agente. Esta é a configuração por defeito que poderá posteriormente servir como uma base para comparações de desempenho.



*Ilustração 10. Exemplo do Interface de conversação simples*

Também é possível configurar o interface para exibir avatares capazes de exibir as seis emoções descritas no modelo anteriormente referido [Ekman e Rosenberg 2005]. As emoções representadas pelo avatar são identificadas no diálogo do agente ou do utilizador através da simples deteção de *emoticons* no texto. Estes *emoticons* também podem ser inseridos no diálogo através do interface e encontram-se definidos na base de dados para poderem ser facilmente alterados.



*Ilustração 11 Exemplo do Interface de conversação com Avatares emotivos*

De forma a ser possível visualizar a quantificação das emoções utilizadas na conversação, é possível utilizar um interface que exibe a percentagem de cada emoção durante a sessão.

Para além das diferenças no interface, a identificação das emoções no texto também é mais precisa. Esta informação é obtida utilizando um API [*EmpathyScope*], adaptado de *Synesketch* [<http://www.synesketch.krcadinac.com>] para funcionar alojado em *Google App Engine*, que funciona de forma semelhante ao mecanismo descrito anteriormente [Liu, Lieberman e Selker 2003].



*Ilustração 12 Exemplo do Interface de conversação com Avatares e gráfico de emoções*

Para que o interface reaja de uma forma mais dinâmica à presença de emoções no diálogo é também possível configurar a aplicação para exibir imagens associadas à emoção predominante. Isto é feito recorrendo a um API que dada uma emoção é possível obter imagens ou citações recolhidas em *blogs* ou outras fontes da Internet. [wefeelfine.org] Estas imagens podem são integradas com o interface substituindo os gráficos de emoções descritos anteriormente.



*Ilustração 13 Exemplo do Interface de conversação com Avatares e imagem de emoção.*

### 4.2.4 Recolha e Armazenamento de Informação

Em todos os tipos de sessão é possível utilizar o interface para mostrar agrado ou desagrado. Este mecanismo é uma forma de obter uma resposta imediata e simples sobre o sentimento geral do utilizador.

Para além da possibilidade de *feedback* do utilizador em qualquer parte do diálogo, a ferramenta possibilita a exibição de um inquérito para obter informação mais precisa acerca da opinião do utilizador. Estas questões para além de permitirem quantificar, em uma escala de 1 a 5, a satisfação do utilizador com a ferramenta em geral, servem também para obter informação acerca da influência do último tipo de interface que utilizaram.

Toda a informação relativa à interação entre o utilizador e o agente é guardada em uma base de dados relacional. Desta forma, será possível visualizar e analisar a informação relativa a todas as sessões e possivelmente identificar alguns dos motivos para as reações negativas e positivas. Assim sendo, é necessário guardar toda a informação relativa ao contexto da conversa, ou seja, o estado emocional de ambos os intervenientes juntamente com a última frase da conversação.

Também serão guardadas as informações relativas a sessões entre dois utilizadores humanos, desta forma poderemos encontrar as reações positivas que poderemos adotar e também identificar algumas reações negativas, que poderão ser ou não reproduzidas pela máquina mas que, obviamente devem ser consideradas como inevitáveis, a não ser que se revele fruto do comportamento de um participante humano, visto decorrerem, de facto, da comunicação entre dois humanos.

## 4.3 Conclusão

Foi especificada e implementada com sucesso uma plataforma para avaliar o desempenho de várias funcionalidades utilizadas em conjunto com um agente de conversação.

Esta plataforma permite o registo de novos utilizadores que poderão avaliar a influência das diferentes funcionalidades do agente através de um conjunto de questionários. Para além de ser possível a interação com um agente de conversação a plataforma permite também a conversação com um humano escolhido para a moderação da plataforma. Desta forma será possível comparar a influência das diferentes funcionalidades quando utilizadas em conversação com um humano e com um agente.

A informação fornecida pelos formulários em conjunto a informação providenciada pelo utilizador recorrendo ao interface para demonstrar a sua satisfação ou insatisfação durante a conversa deve permitir confirmar a influência positiva da representação visual do agente e também verificar a importância da representação emocional da conversação.

Apesar de esta plataforma se apresentar com um conjunto de configurações idealizado para testar um conjunto de funcionalidades específico, esta foi desenvolvida de forma a poder ser experimentada com diferentes configurações e permitir a introdução de novas funcionalidades.

## Capítulo 5

# 5. Avaliação experimental

De forma a comprovar a influência dos elementos multimédia que são utilizados na aplicação, a aplicação foi disponibilizada na Internet e foram convidadas várias pessoas, através de redes sociais, para a utilização da ferramenta desenvolvida.

## 5.1 Configuração da plataforma

Tal como foi descrito no capítulo anterior, a plataforma foi desenvolvida de forma a permitir que se especifique um conjunto de configurações tal como o tipo de sessão e a sua duração. Para a configuração da plataforma para uso experimental, foi determinado que o tempo de sessão devia ser pequeno para evitar que um conjunto maior de sessões se tornasse demasiado demorado. Foram escolhidas 6 sessões de 2 minutos, o que permite o teste de um conjunto significativo de funcionalidades com um tempo médio de utilização da ferramenta à volta de 15 minutos.

Mantendo a proposta do Jogo da Imitação de Turing, estas 6 sessões devem ser divididas entre sessões apenas com intervenientes humanos e sessões de conversação entre um utilizador humano e um agente de conversação. Uma das possibilidades para facilitar disponibilidade para a interação entre 2 utilizadores humanos seria permitir a conversação entre os vários utilizadores. Contudo, para evitar a possibilidade de conversações menos adequadas para o pretendido, ficou definido que no caso de conversação entre humanos, um dos intervenientes teria de pertencer a um conjunto de moderadores, definido anteriormente.

No caso de não haver um moderador disponível para a conversação com o utilizador, será utilizado um agente para o substituir. Caso isto suceda, para evitar que as 3 sessões feitas com cada agente sejam idênticas, nas conversações com o novo agente as conversas devem respeitar um tema específico. Na configuração atual da ferramenta o tema escolhido foi desporto.

A configuração da plataforma foi escolhida de forma às alterações no interface serem sempre incrementais. O utilizador é convidado a falar com um humano e depois com um agente na versão mais simples do interface. Em uma segunda fase as duas sessões aconte-

cem com o interface com avatares. As últimas duas conversas acontecem utilizando o interface com imagens associadas às emoções. Foi considerada a possibilidade de utilizar o interface com os gráficos das emoções mas, neste caso, pareceu mais importante manter a curta duração da fase experimental.

## 5.2 Utilização experimental da plataforma

A plataforma esteve disponível como uma página da Internet e foi feita a publicitação do endereço em redes sociais. Ao longo de 3 dias, registaram-se na aplicação 59 utilizadores para experimentar a ferramenta. De acordo com os dados de registo, estes utilizadores dividem-se em 34 homens e 18 mulheres com um conhecimento médio de Inteligência Artificial. Os restantes utilizadores não especificaram o seu sexo durante o registo.

Decorrente da utilização da plataforma, ficaram registados na base de dados os resultados de 132 questionários. Como cada questionário corresponde a uma sessão que foi efetuada, sabemos que decorreram 264 minutos de conversação, dos quais 202 foram entre um humano e um agente de conversação. Decorreram 31 sessões entre 2 humanos, ou seja, os 62 minutos restantes.

Foram contabilizadas 751 frases de conversação e um total de 115 interações a denunciar a identidade do outro interveniente. Na utilização de este mecanismo os utilizadores indicaram que estavam a comunicar com um agente por 79 vezes contra 36 vezes que escolheram humano. O facto de os utilizadores terem escolhido mais vezes que estavam a comunicar com um agente faz sentido pois o maior de sessões foi mesmo com agentes de conversação.

Para sabermos se os utilizadores responderam, por norma, de forma adequada temos de analisar individualmente os 3 tipos de sessão utilizando os dados da utilização do interface e também as respostas dadas aos questionários.

### 5.2.1 Interface de conversação simples

Após quatro minutos a utilizar a ferramenta os utilizadores conversaram com um Agente de Conversação pelo menos uma vez. Considerando que o tipo de interface é o mesmo para ambas as sessões, as perguntas que responderam ao fim de um minuto também, são idênticas.

Podemos dividir os resultados em 3 categorias de acordo com outro interveniente da conversa. Nomeadamente, quando o utilizador interagiu com um moderador humano, com um agente de conversação normal ou com um agente de conversação restrito ao tópico de Inteligência Artificial.

Os resultados recolhidos pelo controlo do interface que permitem ao utilizador indicar se está a comunicar com um Agente de Conversação ou um humano, exibidos na Tabela 7,

comprovam o que era esperado. Isto é, apesar de os utilizadores se terem enganado várias vezes, em todos os cenários, na sua maioria optaram da forma correta.

O dado mais inesperado é talvez o facto de não haver uma identificação mais precisa da presença de um moderador humano. Este facto é atenuado pela quantidade de vezes que o controlo de identificação foi utilizado. Apesar de o interface notificar o utilizador com alguma frequência acerca da existência de este mecanismo, a sua utilização média foi inferior a duas vezes por sessão.

<b>Destinatário</b>	<b>Humano</b>	<b>Agente de Conversação</b>	<b>Total de Sessões</b>
<b>Moderador</b>	11	8	12
<b>Agente normal</b>	10	19	21
<b>Agente restrito</b>	7	24	25

*Tabela 7 Interação com o controlo de identificação, com interface de conversação simples, por tipo de destinatário*

Para complementar esta informação utilizamos os dados exibidos na Tabela 8, que se referem às respostas ao questionário feito no fim de cada sessão de conversação. Os resultados obtidos através do questionário permitem obter mais informação sobre a opinião do utilizador quanto à sua satisfação ao utilizar a plataforma. De acordo com os dados obtidos, o facto de restringir a conversa a um tema, parece influenciar negativamente o desempenho do agente.

No geral os utilizadores conseguiram identificar com sucesso a identidade do outro interveniente. Mesmo assim, as respostas aos questionários indicam que os utilizadores estariam dispostos a utilizar a ferramenta durante mais tempo e que tiveram uma conversa satisfatória com todos os tipos de participante.

	<b>Moderador</b>	<b>Agente normal</b>	<b>Agente restrito</b>
Acha que esteve a conversar com um humano?	2.58	2.35	2.30
Considera que teve uma conversa satisfatória?	2.75	2.70	2.50
Era capaz de utilizar a ferramenta durante 20 minutos em vez de 2?	3.08	2.90	2.83

*Tabela 8 Médias das respostas aos questionários, com interface de conversação simples, por tipo de interveniente.*

Também podemos ter em consideração o total de emoções expressas, tanto através do uso de *emoticons* como no conteúdo da mensagem. Estes dados, representados na Tabela 9, foram divididos de acordo com a emoção e o tipo de interveniente, onde agora podemos incluir o utilizador. Nesta tabela o dado que se destaca, para além do equilíbrio entre o total

das emoções do utilizador e o total das emoções dos restantes intervenientes, é o facto de se verificar uma ocorrência de Raiva superior a Tristeza. Como o utilizador ainda não foi informado que pode utilizar *emoticons* para exprimir sentimentos, podemos assumir que as emoções foram detetadas no conteúdo do texto.

	Felicidade	Tristeza	Repugnância	Raiva	Medo	Surpresa	Total	Sessões
<b>Moderador</b>	11	2	0	1	2	0	<b>16</b>	<i>12</i>
<b>Utilizador</b>	29	14	0	18	6	0	<b>67</b>	58
<b>Agente Normal</b>	7	6	0	4	2	3	<b>22</b>	<i>21</i>
<b>Agente Restrito</b>	16	2	0	7	8	0	<b>33</b>	25

*Tabela 9 Quantidade de emoções utilizadas, com o interface de conversação simples, por tipo de interveniente*

Os dados relativos aos destinatários das emoções também podem ser analisados e, neste caso, interessa-nos saber para quem foram dirigidas as emoções utilizadas pelos utilizadores da ferramenta.

	Felicidade	Tristeza	Repugnância	Raiva	Medo	Surpresa	Total	Sessões
<b>Moderador</b>	9	3	0	3	1	0	<b>16</b>	<i>12</i>
<b>Agente Normal</b>	7	4	0	5	1	0	<b>17</b>	<i>21</i>
<b>Agente Restrito</b>	13	7	0	10	4	0	<b>34</b>	25

*Tabela 10 Quantidade de emoções utilizadas pelo utilizador, com o interface simples, por tipo de destinatário*

Estes dados, exibidos na Tabela 10, permitem verificar que grande parte das emoções utilizadas pelo utilizador foi em sessões de conversação com um agente restrito. Como o tipo de interface utilizado não exhibe ao utilizador nenhuma informação acerca das emoções utilizadas, sabemos que estes resultados são fruto da conversa e foram as respostas dadas pelo agente que provocaram o maior uso de emoções, neste caso maioritariamente negativas.

### 5.2.2 Interface de conversação com Avatares emotivos

Após a utilização da aplicação com o interface de conversação dotado de *avatares* que exibem as várias emoções dos intervenientes, foi possível obter os dados da Tabela 11. Comparando estes dados com os da sessão anterior, descritos na Tabela 7, torna-se possí-



vel comprovar que a utilização de elementos multimédia, que atribuem características mais humanas ao agente, influencia a experiência de interação.

	<b>Humano</b>	<b>Agente de Conversação</b>	<b>Total de Sessões</b>
<b>Moderador</b>	10	5	10
<b>Agente normal</b>	4	12	18
<b>Agente restrito</b>	4	14	15

*Tabela 11 Interação com o controlo de identificação, com interface de conversação com avatares emotivos, por tipo de destinatário*

As respostas do questionário de este tipo de sessão, exibidos na Tabela 12, mostram que, para além dos utilizadores se mostrarem mais suscetíveis a pensar que estava a conversar com um humano, estes estariam com mais disponibilidade para continuar a conversação do que no tipo de sessão anterior.

	<b>Moderador</b>	<b>Agente normal</b>	<b>Agente restrito</b>
Acha que esteve a conversar com um humano?	2.9	2.44	2.67
Considera que teve uma conversa satisfatória?	3.1	2.52	2.50
Era capaz de utilizar a ferramenta durante 20 minutos em vez de 2?	3.3	2.56	2.80
Observou alterações nas emoções dos intervenientes?	2.9	2.78	2.33

*Tabela 12 Médias das respostas aos questionários, com interface de conversação com avatares emotivos, por tipo de interveniente*

De acordo com os dados do questionário também podemos constatar que os utilizadores tiveram perceção das alterações que ocorreram no interface. As perguntas do questionário poderiam ter sido elaboradas de forma a saber mais detalhes sobre a influência da utilização dos avatares mas isso não foi feito como consequência da curta duração de cada sessão.

	Felicidade	Tristeza	Repugnância	Raiva	Medo	Surpresa	Total	Sessões
<b>Moderador</b>	9	2	0	1	1	1	<b>14</b>	<i>10</i>
<b>Utilizador</b>	38	5	4	22	7	5	<b>81</b>	<i>43</i>
<b>Agente Normal</b>	15	3	0	4	3	0	<b>25</b>	<i>18</i>
<b>Agente Restrito</b>	22	0	0	3	4	1	<b>30</b>	<i>15</i>

*Tabela 13 Quantidade de emoções utilizadas, com o interface de conversação com avatares emotivos, por tipo de interveniente*

Em este tipo de interface, também foram categorizadas as emoções utilizadas pelos vários participantes, tal como é possível ver na Tabela 13. Em este tipo de sessão, tal como nas sessões com o interface de conversação simples, é possível observar que a seguir à Felicidade, a Raiva foi a emoção mais utilizada no discurso dos utilizadores.

Como a ferramenta faz a distinção entre as origens da emoção, sabemos que em este tipo de sessão houve um aumento bastante significativo da utilização de *emoticons*, sendo este mecanismo responsável por 34 das 71 vezes em que o utilizador exprimiu emoções em este tipo de sessão.

Tal como na sessão anterior, interessa-nos saber quem foram os destinatários das emoções utilizadas pelos utilizadores para podermos relacionar essa informação com os dados obtidos anteriormente. Essa informação, representada na Tabela 14, permite concluir que mesmo mudando o tipo de interface, o número de emoções expressas pelo utilizador, se mantém independente do outro elemento da conversação.

	Felicidade	Tristeza	Repugnância	Raiva	Medo	Surpresa	Total	Sessões
<b>Moderador</b>	9	1	0	6	1	4	<b>21</b>	<i>10</i>
<b>Agente Normal</b>	10	3	1	9	5	1	<b>29</b>	<i>18</i>
<b>Agente Restrito</b>	19	1	3	7	1	0	<b>31</b>	<i>15</i>

*Tabela 14 Quantidade de emoções utilizadas pelo utilizador, com o interface com avatares emotivos, por tipo de destinatário*

### 5.2.3 Interface de conversação com Avatares emotivos e imagem de emoção.

Com a adição de um elemento multimédia dinâmico que complementa a informação emocional, seria de esperar uma melhoria no desempenho da aplicação.

Apesar de se verificado um equilíbrio entre as identificações corretas e erradas, recorrendo ao interface, em todos os tipos de sessão, devido ao baixo número de interações, não é possível obter conclusões quanto à eficiência do tipo de interface. Apenas no Agente res-

trito foi possível observar um número de interações mais significativo. Este número também resulta da diminuição de utilizadores entre as várias sessões.

	<b>Humano</b>	<b>Agente de Conversação</b>	<b>Total de Sessões</b>
<b>Moderador</b>	3	4	9
<b>Agente normal</b>	3	3	11
<b>Agente restrito</b>	8	7	11

*Tabela 15 Interação com o controlo de identificação, com interface de conversação com avatares emotivos e imagem de emoção, por tipo de destinatário*

Este valor teria de ser confirmado pela resposta fornecida pelos utilizadores, tal como representado na Tabela 16. As respostas ao questionário evidenciam um resultado menos positivo que parece resultar de erros de decisão na configuração a utilizar na fase experimental. A quebra de satisfação dos utilizadores, quando questionados acerca da última conversação sem restrições, evidencia a saturação resultante do tempo excessivo a conversar com 2 Agentes de Conversação.

Quanto aos utilizadores, que estiveram a conversar com um humano, a melhoria significativa das respostas quanto à satisfação e possibilidade de utilizar a plataforma durante 20 minuto, demonstra a eficiência das alterações na plataforma.

Um valor que se revela interessante é o facto de os utilizadores terem considerado que as emoções utilizadas se traduziram de forma correta em imagens. Este dado serve de validação para o mecanismo de identificação de emoções utilizado.

	<b>Moderador</b>	<b>Agente normal</b>	<b>Agente restrito</b>
Acha que esteve a conversar com um humano?	3.11	1.9	2.18
Considera que teve uma conversa satisfatória?	3.11	2.27	2.50
Era capaz de utilizar a ferramenta durante 20 minutos em vez de 2?	3.44	2.00	2.36
Observou alterações no interface e nas emoções dos intervenientes?	3.22	1.9	3
Considera que as imagens exibidas correspondem às emoções utilizadas?	3.11	2.45	2.63

*Tabela 16 Médias das respostas aos questionários, com interface de conversação com avatares emotivos e imagem de emoção, por tipo de interveniente*

Quanto à quantidade de emoções utilizadas em este tipo de interface, esperava-se um aumento devido aos novos elementos visuais que reforçam o estado emocional dos participantes. Esperava-se que o utilizador utilizasse os avatares e um discurso mais emotivo para poder observar as alterações no interface.

Os dados da Tabela 17, que mostram quais foram as emoções utilizadas pelos vários intervenientes, em conjunto com a Tabela 9 e Tabela 13, que se referem aos mesmos dados nos outros tipo de sessão, permitem concluir que a Felicidade e a Raiva, são as emoções mais utilizadas por todos os tipos de interveniente.

	Felicidade	Tristeza	Repugnância	Raiva	Medo	Surpresa	Total	Sessões
<b>Moderador</b>	6	1	0	0	1	0	<b>8</b>	9
<b>Utilizador</b>	33	6	2	14	1	8	<b>64</b>	31
<b>Agente Normal</b>	14	2	0	5	0	0	<b>21</b>	11
<b>Agente Restrito</b>	10	2	0	2	0	1	<b>15</b>	11

*Tabela 17 Quantidade de emoções utilizadas, com o interface de conversação com avatares emotivos e imagem de emoção, por tipo de interveniente*

Tal como nas sessões anteriores, interessa-nos saber mais sobre as emoções utilizadas no discurso dos utilizadores. Os dados na Tabela 18, que mostram os destinatários das várias emoções do utilizador, evidenciam equilíbrio na forma como se distribui entre os vários tipos de destinatário. Isto é, não podemos concluir que exista alguma diferença na forma como o utilizador comunicou com os vários tipos de interveniente.

	Felicidade	Tristeza	Repugnância	Raiva	Medo	Surpresa	Total	Sessões
<b>Moderador</b>	13	2	0	4	0	0	<b>19</b>	9
<b>Agente Normal</b>	9	3	1	5	1	1	<b>20</b>	11
<b>Agente Restrito</b>	11	1	1	5	0	0	<b>18</b>	11

*Tabela 18 Quantidade de emoções utilizadas pelo utilizador, com o interface com avatares emotivos e imagem de emoção, por tipo de destinatário*

Curiosamente, houve um decréscimo na utilização dos *emoticons* para exprimir emoções. Em este tipo de sessão, das 57 vezes que o utilizador comunicou uma emoção, apenas 22 foram através de *emoticons*.

## 5.3 Conclusão

Depois da análise dos dados recolhidos com a aplicação desenvolvida, foi possível verificar a influência positiva de elementos multimédia quando utilizados em conjunto com tecnologia de Agentes de Conversação. Apesar de não ter sido possível verificar um aumento de satisfação dos utilizadores proporcional aos elementos adicionados ao interface, foi possível manter os níveis de satisfação do utilizador como positivos até uma fase tardia da aplicação.

Quanto à utilização de emoções, é possível observar uma relação entre a forma como o utilizador as exprime e o tipo de interface. A utilização de avatares para exprimir emoções, melhora claramente a experiência de utilização e incentiva o utilizador a ser mais expressivo.

Não foi possível estabelecer uma relação entre o tipo de interveniente e as emoções utilizadas pelo utilizador. De acordo com os dados obtidos, as emoções utilizadas com mais frequência pelo utilizador, isto é a Raiva e a Felicidade, são independentes do tipo de interface e de destinatário.

Infelizmente, devido à configuração definida para a utilização da plataforma, não foi possível obter dados que permitam comparar os vários tipos de interface em circunstâncias exatamente idênticas. Isto acontece porque quando o utilizador é convidado a utilizar o último tipo de interface, já decorreram 8 minutos de utilização. No caso de o utilizador estar a interagir com um humano, este dado não se revela muito influente, mas no caso de serem 2 agentes de conversação, este fator revela-se como negativo na experiência de utilização.

Por outro lado, como a ferramenta foi desenvolvida de forma a permitir a configuração da ordem e duração das sessões, obter dados de utilização da ferramenta utilizada exclusivamente com o interface mais completo, tornou-se uma tarefa bastante simples.

## Capítulo 6

# 6. Conclusões e Perspetivas de Desenvolvimento

Neste trabalho, foram identificados alguns aspetos importantes relacionados com desenvolvimento da tecnologia de Agentes de Comunicação. Recorrendo a estudos anteriores foi possível identificar características importantes para o sucesso dos agentes em diferentes ambientes virtuais. Desta forma, foi possível implementar uma plataforma que permitisse configurar e testar diferentes configurações de ambientes de conversação entre humanos e máquinas. A escolha da tecnologia a utilizar para a implementação da plataforma foi efetuada com o intuito de permitir a facilidade de utilização de elementos multimédia.

Através da plataforma desenvolvida, foi possível efetuar a recolha de dados em várias sessões de teste, recorrendo à utilização por um conjunto de pessoas convidadas para o efeito. Isto é, cada utilizador participará em uma conversação com um Agente de Conversação e outro utilizador, sendo a ordem aleatória. Através da análise da opinião manifestada por cada utilizador, recorrendo ao interface do cliente de conversação, durante ambas as sessões, foi possível comparar o sucesso das diferentes configurações. Desta forma foi possível confirmar os resultados de trabalhos anteriores que enumeram algumas características que deve ser consideradas na utilização de Agentes de Conversação e também identificar outras alterações que podem ser efetuadas tanto na representação do Agente de Conversação como no cliente de conversação que tiveram uma influência positiva no desempenho durante as sessões de teste.

Como desenvolvimento futuro, para além da identificação de novas características que melhorem o desempenho da plataforma, será possível disponibilizar um cliente de conversação que permita a alteração dos seus componentes multimédia. Desta forma será possível a qualquer utilizador, testar a influência de diversos elementos multimédia que foram idealizados por si. Outra forma de desenvolver o trabalho realizado seria através da disponibilização de toda a plataforma com a finalidade de esta poder vir a ser posteriormente melhorada para utilização em trabalhos futuros que incidam sobre o mesmo tema.

## Referências Bibliográficas

- Turing, A.M. (1950). *Computing machinery and intelligence* em *Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer*. Springer, 2009, p.23-65.
- Epstein, Roberts, Gary e Beber, Grace (2008). *Parsing the Turing Test: Philosophical and Methodological Issues*. Springer, 2009.
- Cooper, S. Barry (2004) *The incomputable Alan Turing*. Universidade de Manchester, 5 Junho, 2004.
- Weizenbaum, J. (1966). *ELIZA - A Computer Program For the Study of Natural Language Communication Between Man and Machine* em *Communications of the ACM* Vol. 9, Nº. 1, p.36-45.
- Alice, *Alice Artificial Intelligence Foundation*. Acedido em: 25, Jan. 2010 em: <http://alicebot.blogspot.com/>
- Loebner Prize, *The Loebner Prize in Artificial Intelligence "The First Turing Test"*. Acedido em: 25, Jan. 2010 em: <http://loebner.net/Prize/loebner-prize.html>
- Wallace, Richard, *The Anatomy of A.L.I.C.E.* em *Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer*. Springer, 2009, p.181-210.
- Colby, K. (1975) *Artificial Paranoia: A Computer Simulation of Paranoid Process* em *Artificial Intelligence* Vol. 2, Nº. 1, 1971, p.1-25
- Mauldin, Michael L. (1994) *CHATTERBOTS, TINYMUDS, and the Turing Test: Entering the Loebner Prize Competition* em *Proc. National Conference on Artificial Intelligence AAAI*, Seattle, WA, AAAI Press, The MIT Press, 1994, p.16-21.
- Güzeldere, Güven e Franchi, Stefano (1995) *dialogues with colorful personalities of early A.I.* em *Constructions of the Mind*, Volume 4, Número 2
- Shieber, Stuart M. (1993) *Lessons from a Restricted Turing Test*. *Communications of the ACM*, Vol. 37, Nº. 6, 1994, p.70-78.
- Wilcox, Bruce (2008) *Beyond AIML: Chatbots 102*. Artigo acedido em: 25, Jan. 2010 em: [http://www.gamasutra.com/view/feature/3761/beyond\\_aiml\\_chatbots\\_102.php](http://www.gamasutra.com/view/feature/3761/beyond_aiml_chatbots_102.php)
- Wilcox, Bruce (2011) *Suzette, the Most Human Computer*. Artigo acedido em: 25, Jan. 2010 em: [http://www.gamasutra.com/blogs/BruceWilcox/20110622/7840/Suzette\\_the\\_Most\\_Human\\_Computer.php](http://www.gamasutra.com/blogs/BruceWilcox/20110622/7840/Suzette_the_Most_Human_Computer.php)

- Schumaker, Robert P. e Chen, Hsinchun (2006) *Leveraging Question Answer technology to address terrorism inquiry* em *Decision Support Systems*, Vol. 43, Nº. 4 , 2007, p.1419-1430.
- Schumaker, Robert P.; Ginsburg, Mark; Chen, Hsinchun e Liu, Ying (2006) *Evaluating Mass Knowledge Acquisition using the ALICE Chatterbot: The AZ-ALICE Dialog System* em *International Journal of Man-Machine Studies*, Vol. 64, Nº. 11, 2006, p.1132-1140.
- Schumaker, Robert P.; Ginsburg, Mark; Chen, Hsinchun e Liu, Ying (2007) *An Evaluation of the Chat and Knowledge Delivery Components of a Low-Level Dialog System: The AZ-ALICE Experiment* em *Decision Support Systems*, Vol. 42, Nº. 44, 2007 , p.2236-2246.
- Schumaker, Robert P. e Chen, Hsinchun (2010) *Interaction Analysis of the ALICE Chatterbot: A Two-Study investigation of Dialog and Domain Questioning* em *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, Vol. 40, Nº. 1, 2010, p. 40-51.
- Han, S. e Kim, Y. (2001) *Intelligent Dialogue System for Plane Euclidean Geometry Learning*, apresentado em *International Conference on Computers in Education*, Seul, Coreia, 2001.
- Moore, R. e Gibbs, G. (2002) *Emile: Using a chatbot conversation to enhance the learning of social theory*. Univ. of Huddersfield, Huddersfield, Inglaterra 2002.
- Louwerse, M.; Graesser, A. e Olney, A. (2002) *Good Computational Manners: Mixed-Initiative Dialog in Conversational Agents* em *Papers from the 2002 Fall Symposium, Technical Report FS-02-02*, E. f. H.-C. W. C. Miller, Ed. Menlo Park, CA: AAAI Press, 2002, p.71-76.
- Nass, C.; Steuer, J. e Tauber, E. (1994) *Computers are social actors*, em *Proceedings of CHI '94*, Boston, MA, Abril, 1994, p.72-78.
- Ekman, Paul e Rosenberg, Erika L. (2005), *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System*. Oxford University Press, 1998
- Hadwin, F.H., Winne, P.H. e Nesbit, J.C. (2005). *Roles for software technologies in advancing research and theory in educational psychology* em *British Journal of Educational Psychology*, 75,p. 1-24.
- Lester, J.C., Converse, S.A., Kahler, S.E., Barlow, S.T., Stone, B.A., & Bhogal, R.S. (1997). *The persona effect: Affective impact of animated pedagogical agents* em *Human Factors in Computing Systems: CHI'97 Conference Proceedings*, Nova Iorque, ACM Press.
- Schumaker, Robert P. e Chen, Hsinchun (2010) *Interaction Analysis of the ALICE Chatterbot: A Two-Study Investigation of Dialog and Domain Questioning* em *Systems*,



em *IEEE Transactions on Man and Cybernetics, Part A: Systems and Humans*, Vol 40, Nº 1, p.40-51

Thalmann, N. Magnenat e Kshirsagar, S. (2002) A Multilayer Personality Model. *ACM Symposium on Smart Graphics*, Junho, 2002

Vinciarelli, Alessandro; Pantic, Maja e Bourlard, Hervé (2009) Social signal processing: Survey of an emerging domain em *Image and Vision Computing*, Vol. 27, Nº. 12  
Newton, MA, EUA, Butterworth-Heinemann, Novembro, 2009, p.1743-1759.

Ortony, A.; Clore, GL e Collins, A. (1988) *The cognitive structure of emotions*.  
Cambridge University Press, Maio, 1990

Steunebrink, Bas R; Dastani, Mehdi e Meyer, John-Jules Ch. (2009) The OCC Model Revisited em 4th Workshop on Emotion and Computing, Paderborn, Alemanha, 2009.

Ball, Gene e Breese, Jack (1998) *Emotion and Personality in a Conversational Character* em *Embodied Conversational Character*. MIT Press, 2000, p. 83-84 e 119-121

Thalmann, N. Magnenat; Kim, H. ; Egges, A.; Garchery, S. (2005) *Believability and Interaction in Virtual Worlds* em *MMIEEE Computer Society*, 2005 , p.2-9.

Jabber Software Foundation, *Extensible Messaging and Presence Protocol (XMPP): Core*. Acedido em: 26, Jan. 2010 em: <http://xmpp.org/rfcs/rfc3920.html>

Jabber Software Foundation, *Extensible Messaging and Presence Protocol (XMPP): Instant Messaging and Presence*. Acedido em 26, Jan. 2010 em:  
<http://xmpp.org/rfcs/rfc3921.html>

Jabber Software Foundation, *XMPP Servers*. Acedido em 10, Fev. 2010 em:  
<http://xmpp.org/software/servers.shtml>

Openfire XMPP Server, Ignite Real Time. Acedido em 10, Fev. 2010 em:  
<http://www.igniterealtime.org/projects/openfire/index.jsp>

Xmpp ChatBot, SourceForge. Acedido em 10, Fev. 2010 em:  
<http://xmppchatbot.sourceforge.net/>

Pandorabots. Acedido em 28, Julho 2010 em: <http://www.pandorabots.com>

WeFeelFine. Acedido em 28, Julho 2010 em: <http://wefeelfine.org>

Synesketech. Acedido em 28, Julho 2010 em: <http://www.synesketech.krcadinac.com/>

EmpathyScope Github Acedido em 28, Julho 2010 em:  
[https://github.com/dguttman/empathy\\_scope\\_appengine](https://github.com/dguttman/empathy_scope_appengine)

Lin Shi, Zhiliang Wang, Zhigang Li (2007). *Application and Research on Affection Model Based on Bayesian Network* em *ICEC*, Vol. 4740, Springer, 2007, p.441-444.

- Siska Fitrianie, Pascal Wiggers, Leon J.M. Rothkrantz (2003). *A Multi-modal Eliza Using Natural Language Processing and Emotion Recognition* em *TSD*, Vol. 2807, Springer, 2003, p.394-399.
- Cho, Alvin K. (2007). *Emotional and Domain Concept Enhancements to Alicebot* em *Proceedings of the 2007 Intl. Conference on Artificial Intelligence*, Las Vegas, Junho, 2007
- Vrajitoru, Dana (2006) *NPCs and Chatterbots with Personality and Emotional Response*. *CIG*, IEEE, 2006, p.142-147
- Liu, Hugo; Lieberman, Henry; Selker, Ted (2003) *A Model of Textual Affect Sensing using Real-World Knowledge* em *Proceedings of the 2003 International Conference on Intelligent User Interfaces*, Miami, FL, Janeiro 2003, ACM 2003, p.125-132
- Adjamir M. Galvão, Flávia A. Barros, André M. M. Neves, Geber L. Ramalho (2004). *Adding Personality to Chatterbots Using the Persona-AIML Architecture* em *IBERAMIA*, Vol. 3315, Springer, 2004, p. 963-973.
- Adrian Williamson, Robert Ward (2006). *Emotion in interactive systems: Applying transactional analysis* em *Personal and Ubiquitous Computing*, Vol. 3, Nº. 3, 1999, p. 123-131
- Koda, Tomoko e Maes, Pattie (1996). *Agents with Faces: The Effect of Personification* em *5th IEEE International Workshop on Robot and Human Communication*, 1996
- Beun, Robbert-Jan; Vos, Eveliene de; Witteman, Cilia (2003) *Embodied Conversational Agents: Effects on Memory Performance and Anthropomorphisation* em *IVA*, Vol. 2792, Springer, 2003, p.315-319.
- King, William Joseph e Ohya, Jun (1996) *The Representation of Agents Antropomorphism, Agency and Intelligence* em *Proc. CHI*, 1996
- Xiao, Jun; Stasko, John e Catrambone, Richard (2002) *Embodied Conversational Agents as a UI Paradigm A Framework for Evaluation* em *Proceedings of AAMAS*, 2002
- Cassell, Justine (2000) *More Than Just Another Pretty Face: Embodied Conversational Interface Agents* em *Communications of the ACM*, Vol. 43, Nº. 4, 2000, p. 70-78.
- Zhe, Xu e Boucouvalas, Anthony (2006) *Text-to-Emotion Engine for Real Time Internet Communication* em *International Symposium on CSNDSP*, Staffordshire University, Julho, 2002, p.164-168.
- Graesser, A.C.; Moreno, K.; Marineau, J.; Adcock, A.; Olney, A. e Person, N. (2003). *AutoTutor improves deep learning of computer literacy: Is it the dialog or the talking head?* em *Proceedings of Artificial Intelligence in Education*

Graesser, A.C.; Person, Natalie K.; Harter, Derek e *The Tutoring Research Group* (2001) *Teaching Tactics and Dialog in AutoTutor* em *AutoTutor International Journal of Artificial Intelligence in Education*, 12, p. 257–279.

Person, N.K.; Graesser, A.C.; Kreuz, R.J.; Pomeroy, V. e *The Tutoring Research Group* (2001) *Simulating Human Tutor Dialog Moves in AutoTutor* em *International Journal of Artificial Intelligence in Education* (2001), 12, 23-39

Allen, J. (1999). *Mixed-initiative interaction* em *IEEE Intelligent Systems*, 1999, p.14-23,

Yang, Hui; Chua, Tat-Seng; Wang, Shuguang; Koh, Chun-Keat (2003) *Structured Use of External Knowledge for Event-based Open Domain Question Answering* em *ACM SIGIR 2003*

*“young woman: Men are all alike.*

*eliza: In what way?*

*young woman: They're always bugging us about something specific or other.*

*eliza: Can you think of a specific example?*

*young woman: Well, my boyfriend made me come here.*

*eliza: Your boyfriend made you come here?*

*young woman: He says I'm depressed much of the time.*

*eliza: I am sorry to hear you are depressed.*

*young woman: It's true. I'm unhappy.*

*eliza: Do you think coming here will help you not to be unhappy?*

*young woman: I need some help, that much seems certain.*

*eliza: What would it mean to you if you got some help?*

*young woman: Perhaps I could learn to get along with my mother.*

*eliza: Tell me more about your family.*

*young woman: My mother takes care of me.*

*eliza: Who else in your family takes care of you?*

*young woman: My father.*

*eliza: Your father.*

*young woman: You are like my father in some ways.*

*eliza: What resemblance do you see?*

*young woman: You are not very aggressive but I think you don't want me to notice that.*

*eliza: What makes you think I am not very aggressive?*

*young woman: You don't argue with me.*

*eliza: Why do you think I don't argue with you?*

*young woman: You are afraid of me.*

*eliza: Does it please you to believe I am afraid of you?*

*young woman: My father is afraid of everybody.*

*eliza: What else comes to mind when you think of your father?*

*young woman: Bullies.*

*eliza: Does that have anything to do with the fact that your boyfriend made you come here?"*

**Anexo 1.**

*"How do you do. Please state your problem.  
>> I spend too much time with computers.  
Why do you mention computers ?  
>> You are one.  
Does it please you to believe I am one ?  
>> Yes.  
You seem to be quite positive.  
>> I suppose I am.  
How long have you been ?"*

**Anexo 2.** [Güzeldere e Franchi 1995]

*"01:11:42 Judge 10  
It has been a long day! Now I think I'm ready for some fun. The weather isn't very pleasant today. Do you have any jokes that could lift the clouds?  
01:13:01 Terminal 5  
Everybody talks about the weather but nobody seems to do much about it. Why do you need to tell me how you think you're ready for some fun? Groannnnnnnn....."*

**Anexo 3.** [Shieber Stuart 1993 retirado de *Cambridge Center for Behavioral Studies* 1991]

*"User: Where do I find the web page with Australopithecus?  
Alice: Where do you want to find it?  
User: Bad Answer Alice, time to Learn a new Line  
Alice: How would you answer "Where do I find the web page with Australopithecus?"  
User: I would say: "http://www.ancient skulls.net/australopithecus.html"  
Alice: Thanks, human. Try asking me again.  
User: Where do I find the web page with Australopithecus?  
Alice: http://www.ancient skulls.net/australopithecus.html"*

**Anexo 4.**[*The Elements of AIML Style*]

*“PARRY: I know the mob controls the big rackets.*

*ELIZA: Who are you really talking about?*

*PARRY: Have you ever really thought about the Mafia?*

*ELIZA: Suppose you ever really thought about the Mafia?*

*PARRY: Let's talk about something else.*

*ELIZA: I am not sure I understand you fully.*

*PARRY: You keep asking me the same things, why not change the topic?*

*ELIZA: In what way?”*

**Anexo 5.**[Güzeldere, Franchi 1995]

*“jabberwacky: Are you an alien?*

*user: Yes, a pink alien.*

*jabberwacky: Not quite!*

*user: Yes, I am an alien, and I am pink.*

*jabberwacky: Are you human?*

*user: Half. I am half human half alien.*

*jabberwacky: That's weird.*

*user: Yes. I like being weird.*

*jabberwacky: Well, you are pretty good at it.*

*user: Thanks.*

*jabberwacky: It's not a compliment.*

*user: But I take it as a compliment.*

*jabberwacky: Well done.*

*user: Thanks.”*

**Anexo 6.** [<http://jabberwacky.icogno.com/j2convbycategory>]